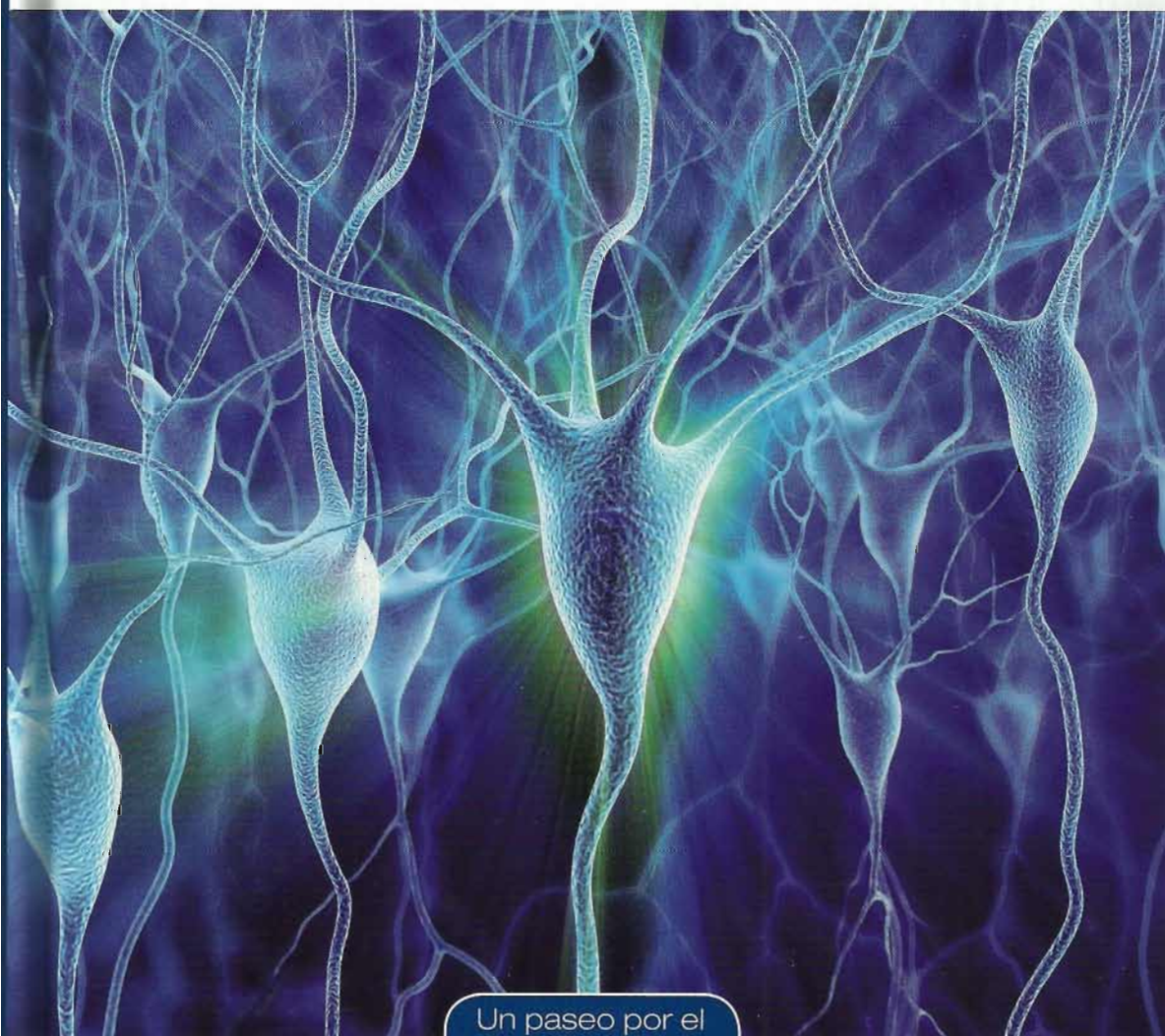


Ciencia y consciencia

La interacción entre mente
y materia



Un paseo por el
COSMOS



The Doctor y La Comunidad

Redigitalización: The Doctor

<http://thedoctorwho1967.blogspot.com.ar/>

<http://el1900.blogspot.com.ar/>

<http://librosrevistasinteresesanexo.blogspot.com.ar/>

<https://labibliotecadeldrmormoreau.blogspot.com/>

Ciencia y consciencia

La interacción entre mente
y materia

RBA

Imagen de cubierta: El cerebro está formado por una enorme red de neuronas conectadas entre sí. Pero ¿es el pensamiento reducible a una serie de impulsos eléctricos? ¿O hay algo más allá de la física que determina nuestra consciencia?

Dirección científica de la colección: Manuel Lozano Leyva

© Eduardo Arroyo Pérez por el texto
© RBA Contenidos Editoriales y Audiovisuales, S.A.U.
© 2016, RBA Coleccionables, S.A.

Realización: EDITEC

Diseño cubierta: Llorenç Martí

Diseño interior: tactilestudio

Infografías: Joan Pejoan

Fotografías: Archivo Pejoan: 26, 27b; George Ioannidis: 93;
Jordgette: 31; Dennis Kunkel/Age Fotostock: 102; Pasioka/SPL/Age
Fotostock: portada; Royal Society: 27a; Stigmatella aurantiaca: 59;
Dr. Tonomura: 29, 32; Vignolini: 27c.

Reservados todos los derechos. Ninguna parte de
esta publicación puede ser reproducida, almacenada
o transmitida por ningún medio sin permiso del editor.

ISBN: 878-84-473-8550-7

Depósito legal: B-0883-2016

Impreso y encuadernado en Rodesa, Villatuerta (Navarra)

Impreso en España - Printed in Spain

SUMARIO

INTRODUCCIÓN	7
CAPÍTULO 1	El problema de la medida 13
CAPÍTULO 2	Tiempo y consciencia en la relatividad de Einstein . . . 61
CAPÍTULO 3	Física y libre albedrío 83
CAPÍTULO 4	Cómo simular la mente 115
LECTURAS RECOMENDADAS	155
ÍNDICE	157

En el mundo del conocimiento hay una tendencia a separar y clasificar en áreas bien delimitadas todo lo que sabemos. Mientras que los psicólogos se ocupan de la consciencia, los físicos hacen lo propio con las partículas. Y los filósofos... pues se ocupan de todo un poco, para irritación de algunos. En general, los expertos de un área se mantienen en ella y son criticados, a veces con virulencia, si intentan aplicar su conocimiento más allá de los límites de su disciplina.

Esta separación en áreas, aunque sea artificial, resulta beneficiosa en muchos casos, porque delimitar bien un problema suele ser el primer paso para su solución. Ciertamente, la física posee capacidad predictiva gracias a que se centra en unos pocos aspectos cuantitativos de los sistemas que estudia, dejando el resto para otras ciencias o saberes. De la misma forma, la psicología deja el estudio del cerebro a la neurología y se ocupa de la interacción entre estados mentales. En ambos casos, sacrificamos información para ganar comprensión: un modelo solo es útil si contiene menos datos que la realidad. Por tanto, un físico tiene que asumir que su humor no tiene ninguna influencia en los resultados de sus experimentos. De otra forma, le sería imposible seguir haciendo ciencia.

Sin embargo, el siglo xx reveló que, a veces, la separación entre áreas no es posible. La física, el paradigma de la ciencia objetiva, fue atacada por un torrente de subjetividad: los físicos descubrieron que un investigador tiene influencia en lo que mide por el solo hecho de estarlo midiendo. Cualquier intento de ignorar esa realidad lleva a realizar predicciones incorrectas. El experimentador, al ser parte del propio sistema que está estudiando, tiene que ser tomado en consideración, lo que causa un sinnúmero de problemas técnicos, filosóficos y matemáticos.

Las dos grandes teorías del siglo pasado, la mecánica cuántica y la relatividad, comparten que la irrupción del observador es una variable que, lejos de poder ignorarse, juega un papel central. Quién mide qué y cómo lo hace es tan importante como lo que se está midiendo. Cuestiones como la consciencia, consideradas tradicionalmente sospechosas por la física, tienen que ser tenidas en cuenta, resquebrajando el edificio de objetividad absoluta que la ciencia intenta construir.

Hoy, los interrogantes planteados por la física del siglo xx siguen abiertos. La mecánica cuántica tiene varias interpretaciones que, a pesar de hacer las mismas predicciones, describen realidades radicalmente distintas. La validez de estas interpretaciones sigue siendo objeto de agria polémica entre los físicos, sin que parezca que haya un claro vencedor. La comunidad está dividida. Unos se han atrincherado en el bando realista, que cree en una realidad objetiva y externa a nosotros, y otros en el subjetivista, convencidos de que el mundo solo toma existencia cuando se mide. Ambos bandos se acusan de no haber entendido las lecciones del siglo xx. ¿Es la consciencia lo único real o es tan solo una ilusión creada por el mundo?

Si la comunidad física se vio alterada por una serie de problemas surgidos del infierno metafísico, la filosofía tampoco salió libre del encuentro: los nuevos descubrimientos en física hacían afirmaciones muy claras sobre el mundo microscópico que no podían ignorarse en el estudio de la consciencia y del libre albedrío. Hoy en día, el modelo estándar es capaz de predecir todos los fenómenos conocidos de nuestra experiencia cotidiana, lo que resulta uno de los mayores logros de la humanidad hasta

la fecha. La física ha salido de su ámbito y se ha colado en los despachos de la filosofía.

Los éxitos de la física en la predicción de fenómenos subatómicos han dado alas al reduccionismo, la creencia que afirma que todo lo que sucede en el universo puede reducirse a la interacción entre partículas elementales. Hasta la fecha, no se ha encontrado ningún comportamiento de la materia que vaya en contra de esta hipótesis: todas y cada una de las partículas que nos encontramos parecen hacer exactamente lo que establecen las leyes de la física, sin excepción.

Sin embargo, el reduccionismo pone nerviosa a mucha gente. Si las personas no somos más que una colección de átomos y moléculas que se comportan de forma mecánica, ¿significa eso que no somos más que autómatas? ¿Qué hay de nuestra libertad para tomar decisiones? ¿Qué hay de nuestra consciencia de estar vivos? ¿De dónde sale todo eso? La ciencia es capaz de dar respuestas a algunas de estas cuestiones, pero no a todas. Por desgracia, no existe tampoco una explicación alternativa: si la consciencia no es generada por la interacción entre nuestras partículas ¿qué la causa? Nadie lo sabe.

En este campo, como en muchos otros, se libra una batalla encarnizada entre ciencia y sentido común. Doscientos años atrás, las leyes de Newton posibilitaron la reconciliación de las predicciones de la física y las intuiciones humanas sobre el mundo: el universo podía ser imaginado como un gran engranaje, con la gravedad como pegamento cósmico. El siglo xx se caracterizó, en cambio, por mostrar los límites de nuestra intuición. La mecánica cuántica dio al traste con la creencia de que las partículas tienen que estar en una posición determinada; la relatividad acabó con nuestra idea del tiempo como algo que fluye. En ambos casos, los científicos que las descubrieron tuvieron que elegir entre confiar en su intuición o en sus datos: eligieron lo último. No todo el mundo les acompañó e, incluso hoy, muchos se resisten a aceptar una serie de consecuencias que no casan con sus prejuicios sobre el universo.

Algo parecido se da con el problema del libre albedrío. Nuestra intuición nos hace creer que tomamos decisiones libres a

cada momento y que son estas decisiones las que determinan nuestro futuro. También nos dice que hay decisiones buenas y malas y que somos responsables de ellas. Por otro lado, la física newtoniana nos asegura que nuestras decisiones son un producto de la interacción entre nuestras partículas y que, aunque nos parezca que lo hacemos libremente, en realidad solo seguimos el dictamen de las leyes. Se trata de una cuestión especialmente difícil, porque en este caso nuestras intuiciones son muy potentes: somos capaces de sentir cómo tomamos una decisión, con lo cual tenemos un conocimiento que se nos antoja directo, tanto como el de observar el color del cielo. Por mucho que nos digan que ese color no existe, resulta difícil de creer. Al fin y al cabo, lo estamos viendo.

El problema de la consciencia es, si cabe, más peliagudo. El libre albedrío es una sensación, pero la consciencia es la condición previa a tener sensaciones. Resulta difícil decir algo preciso sobre la consciencia, aparte del mero hecho de que poseemos una. Sin embargo, tratar de averiguar qué es o de dónde surge por introspección es una tarea imposible, porque la propia consciencia precede a cualquier construcción mental. El conocimiento de que tenemos consciencia resulta aún más difícil de reconciliar con una visión del universo basada en pequeñas partículas que interactúan con reglas establecidas. Si todo lo que hay en el universo son partículas que chocan entre sí, ¿cómo surge la consciencia? ¿Cómo podemos explicar algo como la sensación de ver rojo solo a partir de las interacciones entre las partículas de nuestro cerebro? ¿Qué hay en las ecuaciones de la mecánica cuántica que pueda describir lo que se siente al escuchar una canción? Todas esas son preguntas que podemos hacer a la física, pero para las que la física no tiene, de momento, respuesta alguna. Son también el tipo de preguntas que se usan para argumentar que tiene que haber algo más: algo más allá de la física que sea capaz de explicar ese fenómeno casi mágico que es el universo preguntándose sobre sí mismo.

La física, por definición, no puede abordar el problema de la consciencia. La ciencia se basa en la predicción de resultados experimentales y debe ser, por consiguiente, comunicable. La

experiencia subjetiva es todo lo contrario. Toda teoría científica tiene que ser susceptible de ser falsada: es decir, tiene que poder hacer predicciones concretas cuya validez o invalidez determinarán su suerte. Sin embargo, no hay ningún experimento posible que nos pueda decir si un organismo es consciente. Por supuesto, podemos aceptar pruebas indirectas: por ejemplo, si el organismo nos dice que es consciente, podríamos elegir creerle, pero seguiríamos sin tener una prueba fehaciente. Esto hace que cualquier teoría de la consciencia sea imposible de verificar, lo que inmediatamente la convierte en acientífica. El problema de la consciencia tiene que ser abordado, entonces, por la filosofía, aunque sin perder de vista todo lo que nos enseña la ciencia al respecto.

Si bien es cierto que la física no puede despejar el enigma de la consciencia, también lo es que, sin la física, estamos dando palos de ciego. Abordar el problema de la consciencia ignorando lo que sabemos sobre el mundo microscópico parece una empresa destinada al fracaso: cualquier explicación deberá tener en cuenta a la física, aunque puede que vaya más allá de esta.

En tiempos recientes, otras disciplinas se han convertido en serios candidatos a proporcionar la nueva pieza del rompecabezas: en particular, la inteligencia artificial ha dado pasos de gigante en los últimos diez años, programando ordenadores para realizar tareas hasta hace poco reservadas a las personas. La neurología también ha experimentado grandes avances, sobre todo gracias a nuevas técnicas no invasivas para registrar el comportamiento del cerebro en tiempo real. Curiosamente, esas técnicas se basan en desarrollos realizados en el marco de la física de partículas.

Aunque ciencia y consciencia parecen habitar reinos distintos, los descubrimientos en la física del siglo xx las han unido sin remedio. Hoy en día no es posible comprender una sin la otra: para poder entenderse, el universo debe saber que se observa a sí mismo.

El problema de la medida

Sin duda resulta chocante que el impredecible comportamiento humano pueda estar regulado por las mismas leyes que rigen el universo. Por ello, cuando la mecánica cuántica forzó la revisión del papel de la consciencia, provocó un cambio de paradigma que fue traumático para gran parte de la comunidad física.

«¿Qué es, pues, el tiempo? Si nadie me lo pregunta, lo sé; si quiero explicarlo a quien me lo pide, no lo sé.» Esta cita de san Agustín sobre el paso del tiempo también podría aplicarse a la consciencia, a la que podríamos describir como la única parte del universo a la que, en principio, tenemos acceso. El resto es una inferencia lógica: de nuestro conjunto de sensaciones deducimos que debe de haber, por ejemplo, libros y casas, así como otras personas que comparten nuestra experiencia, es decir, un mundo exterior a nosotros mismos. La consciencia, al ser una parte tan esencial de nuestra existencia, parece existir en un segundo plano, como si se tratara de algo que damos por sentado mientras nos fijamos en lo que sucede en el mundo.

La relación entre consciencia y materia es aún objeto de debate por parte de científicos y filósofos. De hecho, es difícil encontrar un tema sobre el que la comunidad científica esté más dividida. Por un lado, algunos científicos creen que la consciencia es un producto de las leyes de la física y, como tal, puede ser recreado en el laboratorio. De hecho, existen ya intentos de crear consciencias artificiales usando superordenadores. Por otro, hay pensadores que consideran que la consciencia es algo

fundamental y distinto de la materia y que por lo tanto, nunca podrá ser explicada, ni mucho menos simulada, por la ciencia.

¿Y TÚ QUÉ ERES, MONISTA O DUALISTA?

A la creencia de que mente y materia son lo mismo se le suele llamar *monismo* y a la contraria, *dualismo*. En la actualidad la mayoría de científicos podrían clasificarse como monistas, pero no se trata de un debate zanjado en absoluto. Aunque hoy podría parecer que el monismo es una opción más científica y más moderna que el dualismo, lo cierto es que es anterior a este; la introducción del dualismo ayudó a cimentar la revolución científica que detonaron, entre otros, Galileo, Newton y Descartes. Si preguntásemos a una persona de hace 50 000 años si la mente y la materia están separadas, muy probablemente contestaría que no, a juzgar por lo que los antropólogos han descubierto tras estudiar tribus con un nivel similar de desarrollo científico. De hecho, esa persona ni siquiera sería consciente de ser monista, ya que su creencia sería tan arraigada que no habría alternativa. Parece ser que los humanos son monistas «por defecto».

Pero el monismo de los pueblos prehistóricos tenía poco que ver con el monismo de hoy en día. Por ejemplo, en su visión del mundo la lluvia no caía porque la empujase la gravedad, sino porque «quería» caer. Las piedras eran difíciles de transportar porque no se querían mover. En el monismo prehistórico, que se suele llamar *animismo*, no era la mente la que se regía por las leyes del universo, sino que era el universo el que se regía por las leyes de la mente.

En la física de Aristóteles (384-322 a.C.) aún puede verse un resto de ese monismo ancestral. En su sistema, el mundo estaba constituido por cinco elementos, cada uno con su lugar natural. La tierra ocupaba el centro del planeta, seguida por el agua, el aire y el fuego. Finalmente, los astros estaban hechos de otro material llamado éter. Así, las piedras no caían porque las empujase una fuerza externa, sino que lo hacían porque era su tendencia natural moverse hacia el centro de la Tierra. De la mis-

ma forma, la lluvia caía hacia el suelo para dirigirse al océano, siempre por encima de la tierra. Al alterar el estado natural de las cosas, por ejemplo elevando una piedra, se notaba una resistencia que no era más que la tendencia de esa piedra a volver a su lugar natural.

Dotar a los objetos inanimados de voluntad es algo natural para las personas y, por eso, deshacerse de esa intuición fue una tarea que requirió cientos de años. Hubo algunos intentos antes de Aristóteles: Demócrito (460-370 a.C.), siguiendo a su maestro Leucipo, propuso la idea de que el universo estaba constituido por átomos, pequeñas partículas indivisibles en perpetuo movimiento. De hecho, la palabra griega átomo significa literalmente «que no se puede cortar, indivisible». La forma de esos átomos explicaría las propiedades de las diferentes sustancias. La teoría de Demócrito fue un primer intento de descripción *mecanicista* del universo: los átomos «no querían» nada, sino que simplemente se movían según fuerzas que escapaban a su control. A su vez, la teoría de Demócrito era monista, ya que se aplicaba también a las personas. Esto era algo difícil de aceptar: si la gente estaba hecha de átomos que seguían rígidos patrones de movimiento ¿dónde quedaba el espacio para el libre albedrío? Lucrecio (99-55 a.C.), inspirado en la física de Epicuro (341-270 a.C.), propuso la teoría del *Clinamen* (derivado de *clinare* o inclinar) según la cual, de cuando en cuando, los átomos describían movimientos aleatorios y, por lo tanto, impredecibles, lo que explicaría la libertad humana.

La idea de Demócrito fue abandonada durante cientos de años, superada en popularidad por la física aristotélica, más en línea con el sentido común. Al no distinguir entre mente y materia, cualquier teoría mecanicista de la materia se vería con el mismo problema: ¿por qué no opera así la mente? El monismo imperante representaba un obstáculo formidable para el desarrollo de la ciencia, basada en la creencia de que el universo opera según leyes predecibles.

El animismo aristotélico empezó a ser superado con el trabajo del célebre astrónomo italiano Galileo Galilei (1564-1642), quien abandonó las tendencias naturales de Aristóteles para abrazar

la entonces novedosa idea de que el universo se regía por leyes matemáticas, Galileo tuvo muchos problemas con la Iglesia, porque sus tesis entraban en conflicto con ciertos pasajes bíblicos. Además, eliminar la voluntad del comportamiento de la materia representaba una amenaza: si los cuerpos se comportaban según unas normas matemáticas preestablecidas, ¿dónde quedaba Dios?, ¿qué papel jugaba en el devenir cotidiano?, ¿se limitaba tan solo a poner en marcha el universo, como si de un juguete de cuerda se tratase? Un mundo con leyes está exento de milagros. Es más: un mundo donde esas mismas leyes se aplican también a las personas parece ser, también, un mundo sin libre albedrío, que forma la base del bien y del mal.

Es posible que el matemático y filósofo francés René Descartes (1596-1650) fuera muy consciente de los problemas del monismo, porque su filosofía introdujo una solución que lograba conciliar el aspecto mental y material. Según Descartes, el cuerpo era material, pero la mente no; ambos interactuaban a través de la glándula pineal. Descartes veía el cuerpo como un autómata, sujeto a las mismas leyes de la naturaleza que un objeto inanimado. La mente, por otra parte, era inmaterial y no se regía por las leyes físicas, aunque tanto podía ser afectada por el cuerpo como influir en este. De esta forma, era posible estudiar el universo como si fuera un mecanismo, sin tener que renunciar ni a la libertad, ni al alma ni a Dios. Pero a pesar de los esfuerzos de Descartes para congraciarse con la Iglesia, en 1663 el papa incluyó sus escritos en el índice de libros prohibidos.

Ya impuesto el dualismo, la física tuvo las manos libres para centrarse en el comportamiento de la materia sin tener que entrar en temas metafísicos. Mientras tanto, la mente, con sus complicaciones y misterios, quedó relegada en manos de la religión, la filosofía de la mente y la psicología.

LOS AUTÓMATAS DE HUXLEY: EL RETORNO DEL MONISMO

La tregua entre ciencia y religión duró poco. Los constantes éxitos de la ciencia, en especial de la física newtoniana, llevaron a

muchos intelectuales a replantearse el alcance del método científico y de las explicaciones científicas en general. Si fenómenos tan dispares como las nubes o la luz eran explicables en términos mecánicos, ¿por qué no lo iba a ser la mente? Era una progresión lógica. Después de que Descartes considerara que el cuerpo era un mero autómata conectado al alma por la glándula pineal, el siguiente paso fue prescindir del alma y dar por supuesto que el autómata venía con consciencia incorporada, sobre todo tras comprobar que dicha glándula pineal no operaba en el ser humano tal y como decía el filósofo francés. Lo que el dualismo había separado lo volvió a unir el mecanicismo. No obstante, se había pasado de un universo mental a una mente mecánica. Mente y materia volvían a unirse, pero, esta vez, la voz cantante la llevaba la materia.

Sin embargo, considerar que la física puede explicar la consciencia dista mucho de explicar cómo lo hace. Resultó ser un problema mucho más complejo de lo que se creía en un principio. Tanto, que hoy seguimos lejos de encontrar la solución. Los primeros intentos serios llegaron a finales del siglo XVIII y principios del XIX. Las observaciones científicas realizadas hasta la fecha mostraban que explicar el funcionamiento del cuerpo como si se tratara de un mecanismo autómata se correspondía en buena medida con la realidad: el corazón es una bomba que se encarga de que circule la sangre; el hígado, un filtro que elimina toxinas; los pulmones proporcionan el oxígeno necesario para quemar el combustible que llega en forma de comida. Si el cuerpo en su totalidad era un mecanismo, nada hacía suponer que el cerebro fuese diferente. Entonces... ¿era simplemente una máquina que hacía lo que le dictaban las leyes de la física? Y en ese caso, ¿qué papel jugaba la consciencia?, ¿por qué tenemos la sensación de que tomamos decisiones que luego nuestro cuerpo ejecuta? ¿Cómo explica eso la teoría mecánica de la consciencia?

Durante el siglo XIX, varios científicos, entre los que destaca Thomas Henry Huxley (1825-1895), aquel biólogo apodado «el bulldog de Darwin» por su enardecida defensa de la teoría de la evolución, propusieron que la consciencia es un *epifenóme-*

no: algo así como un efecto colateral de la operatividad del cerebro que no tiene causa alguna. Según Huxley, puede parecer que nuestras intenciones son la causa de nuestras acciones, pero esto no es así: en realidad, nuestro cerebro actúa de manera automática, siendo la consciencia un subproducto de su operación. En el esquema de Huxley, la consciencia es similar al pitido de una máquina de vapor: a pesar de que el arranque del tren coincide con este, se trata solo de una coincidencia en el tiempo y no de una relación de causa y efecto. Lo que pone en marcha al tren es el motor, aunque no lo percibimos directamente. De la misma forma, los mecanismos que conducen a la acción de nuestro cuerpo permanecen ocultos, aunque tengamos la ilusión de que somos nosotros los que los determinamos.

El nuevo monismo materialista relegaba a la consciencia a un papel secundario, como si no fuera más que un mero espectador de un mundo mecánico. Esa nueva visión no tuvo, sin embargo, efecto en el desarrollo de la física como ciencia: tanto si uno era dualista como monista, la consciencia permanecía como un ente pasivo, separado del universo medible y, por tanto, sin consecuencia alguna para el comportamiento de la materia. Podría hablarse de un monismo desconectado: incluso aceptando que mente y materia son indisociables, no había razón para mezclarlos. Ese *status quo* estaba, sin embargo, a punto de sufrir un golpe mortal.

CONSCIENCIA Y REALIDAD CUÁNTICA

El principio del siglo xx fue un periodo convulso para la física, que vio desmoronarse el edificio de la mecánica clásica, el cual fue reemplazado por las dos teorías que dominan en la actualidad: la relatividad de Einstein y la mecánica cuántica. Estas dos teorías obligaron a los físicos a replantearse todo lo que daban por sentado sobre el espacio y el tiempo, así como la naturaleza de las partículas que constituyen el mundo. La mecánica cuántica les forzó, además, a abandonar la idea de una consciencia pasiva y separada del resto del universo.

A principios del siglo xx todo el mundo daba por sentado que la realidad era la que era, independientemente de quién la percibiera o de cómo lo hiciera. Al fin y al cabo, cuando alguien dice que observa algo que nadie más ve, se da por supuesto que sufre algún trastorno o que está bajo los efectos de alguna droga. También imperaba la idea de que el mundo se comportaba de la misma forma tanto si era observado como si no: la Luna sigue ahí aunque no la estemos mirando.

Estas dos suposiciones estaban ancladas en millones de observaciones realizadas a lo largo de cientos de años: cuando volvemos a casa después de un día de trabajo, todo sigue como lo hemos dejado y nuestra familia puede atestiguar que nada ha cambiado mientras hemos estado fuera. Por supuesto, podría pasar que el universo se desvaneciese en el momento en el que apartamos la vista y volviera a aparecer cuando lo miramos, pero esa posibilidad parece poco plausible, al ser mucho menos simple que asumir lo contrario. La idea de que una partícula cambiase de comportamiento al ser observada hubiera resultado ridícula para un físico del siglo xix. Sin embargo, eso es lo que se vieron obligados a aceptar los físicos del siglo xx.

Las pruebas de que algo no encajaba con la física tradicional fueron acumulándose a lo largo de la primera mitad del siglo xx. En las páginas que siguen no se dará una explicación cronológica, sino que nos centraremos en el llamado «experimento de la doble rendija», uno de los experimentos fundacionales de la física cuántica que pone de manifiesto, mejor que ningún otro, el papel de la consciencia.

Para ello, es conveniente empezar primero hablando de otro experimento relacionado. Se necesita una pistola de balines, una lámina con dos rendijas y una pantalla donde quede marcada la posición de impacto de cada balín, como se muestra en la figura 1.

La pistola se supone anclada a un soporte y solo puede rotar. Una vez establecido el dispositivo experimental, se procede a disparar los balines en direcciones aleatorias, tantos como sea posible. Después de un gran número de disparos, se tendría que obtener un resultado parecido al ilustrado en la figura 2.

Es posible realizar un experimento diferente, usando ondas en lugar de balines. Una onda es la propagación de una perturbación en un medio, como hacen las olas en el agua. El sonido o la luz son otros ejemplos. En este caso supondremos que usamos ondas en el agua, aunque este experimento funcionaría también con sonido o con luz.

FIG. 1

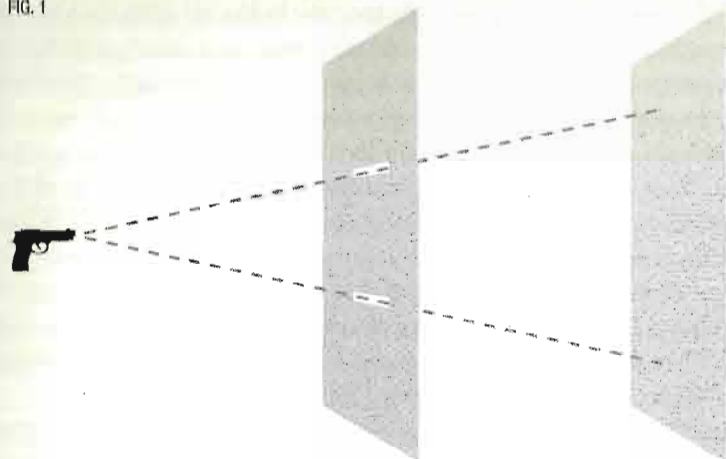


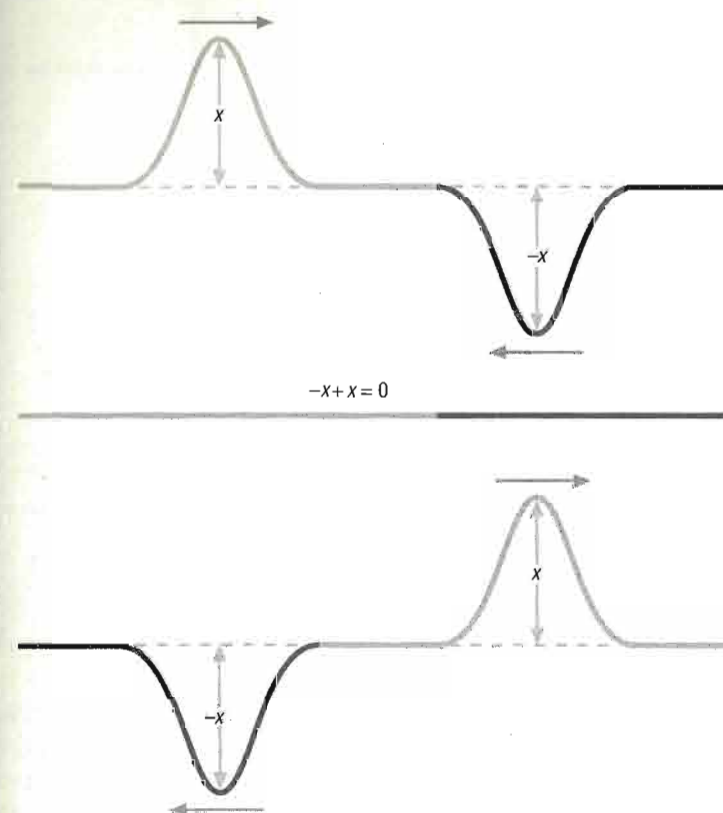
FIG. 2



Las figuras ilustran un tipo de experimento de doble rendija en el que se disparan balines de forma aleatoria. Los impactos se concentran en la parte superior e inferior de la pantalla.

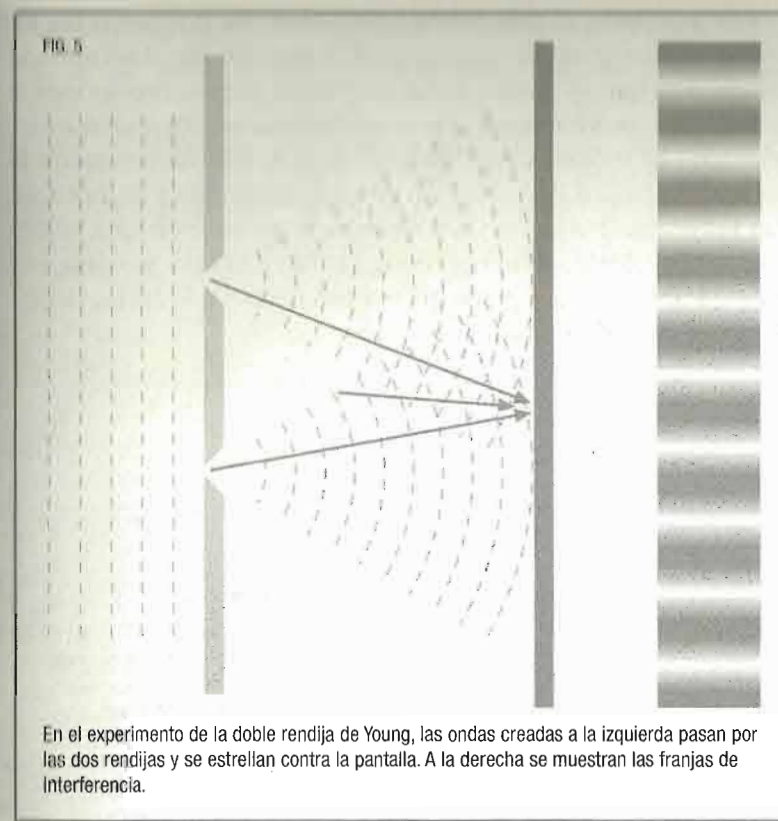
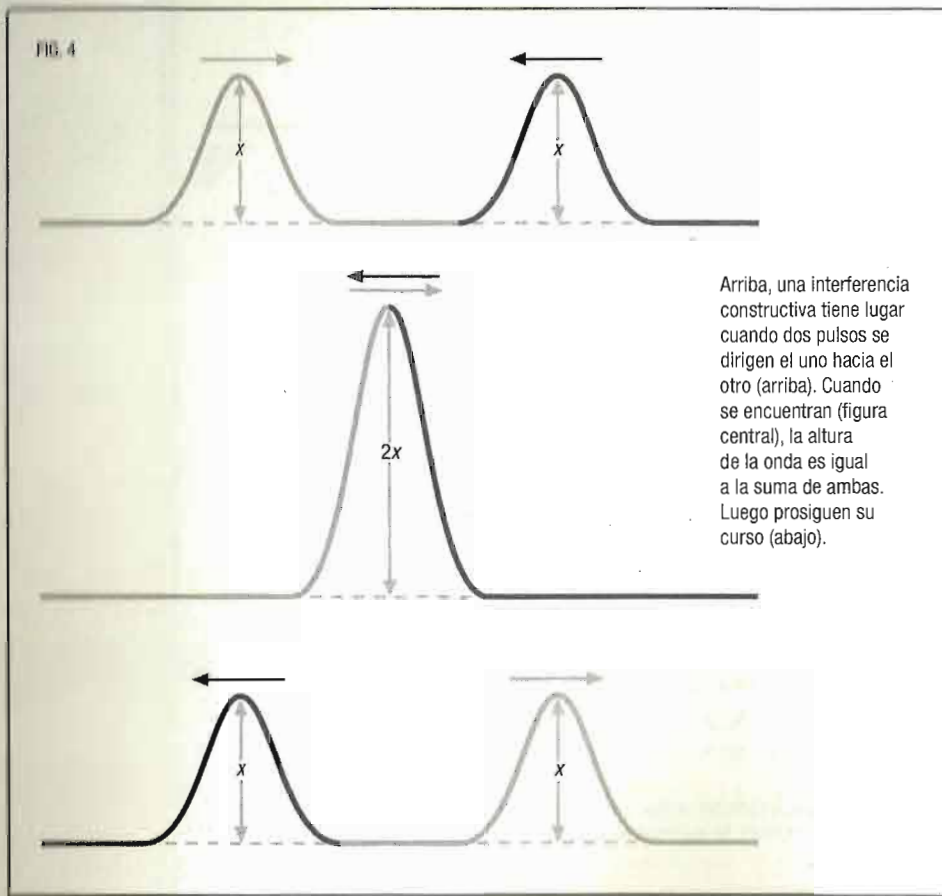
Cuando dos ondas se propagan por un medio, pueden interactuar la una con la otra en un fenómeno denominado *interferencia*. Cuando la parte más alta de una de las ondas (la *cresta*) coincide con la más baja de la otra (el *valle*) las dos perturbaciones se cancelan entre sí, dando lugar a algo llamado *interferencia destructiva*, que se muestra en la figura 3.

FIG. 3



Una interferencia destructiva (arriba) tiene lugar cuando dos pulsos se dirigen el uno hacia el otro y, al encontrarse, se cancelan mutuamente (centro). Luego siguen su curso de nuevo (abajo).

Por otro lado, si dos crestas coinciden, las dos perturbaciones se suman, dando lugar a una onda que alcanza determinada altura, a la que se suele llamar *amplitud*, y que corresponde a la suma de las anteriores. A esto se le llama *interferencia constructiva* (figura 4). La interferencia es un fenómeno común en la naturaleza (véase la ilustración de las págs. 26-27) y se da en todas las ondas. De hecho, la observación de interferencia ayudó al científico inglés Thomas Young (1773-1829) a demostrar a la comunidad científica que la luz era una onda y no una partícula, como se creía desde tiempos de Newton.

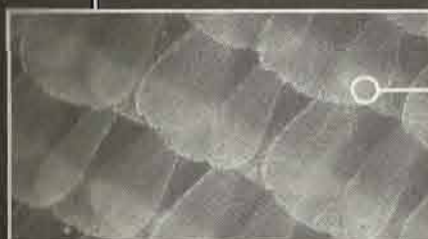


Young es célebre por realizar el experimento de la doble rendija (figura 5) con ondas, muy parecido al anterior, pero en el que la pistola de balines es sustituida por un generador de ondas, que normalmente es una lámina metálica que oscila a un ritmo determinado; la siguiente parte del experimento es prácticamente igual. Se dispone de una barrera con dos aperturas que se suponen pequeñas y de una pantalla donde chocarán las ondas. En este caso, no se mide el número de impactos, sino la intensidad de las ondas que chocan contra la pantalla.

Lo que se detecta en esa pantalla es que en cada punto se juntan dos ondas que han tenido recorridos diferentes, al haber pasado por rendijas distintas. Según la diferencia de sus trayec-

JUEGOS DE LUZ: INTERFERENCIA DE ONDAS EN LA NATURALEZA

La interferencia genera fantásticos matices iridiscentes en mariposas, colibríes, insectos y determinadas plantas. Ciertas mariposas del género *Morpho* presentan un azul espectacular debido a la estructura doble de sus escamas. Los resplandecientes tonos metálicos del escarabajo joya son fruto de una cutícula multicapa que distorsiona la refracción de las ondas de luz, y las llamativas plumas de los colibríes responden no solo a sus pigmentos, sino a la refracción provocada por el complejo «armazón» de melanina de sus plumas. Por su parte, la planta africana *Pollia condensata* brilla gracias a su cutícula, que refleja la luz de forma especular.



MARIPOSAS MORPHO

La estructura de las escamas de estas mariposas provoca que las ondas de luz se sometan a una interferencia constructiva y destructiva. ¿El resultado? Un porte iridiscente que se extiende por toda el ala.

Escamas dentadas.

Lámina base de las escamas alares



ESCARABAJO JOYA

Los élitros de este coleóptero están repletos de pequeñas rugosidades y muescas, lo que afecta a la reflexión. Según la curvatura de la onda, pueden adquirir diversos colores y tonalidades.



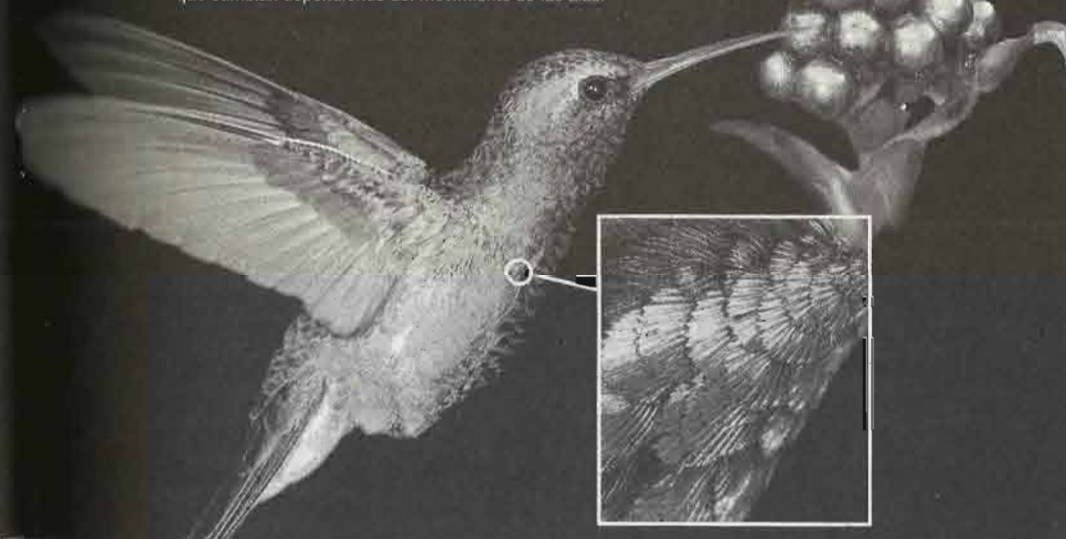
UNA BAYA AFRICANA: POLLIA CONDENSATA

Este fruto tan brillante adopta un aspecto intensamente metalizado debido a que su superficie, compuesta por cuatro capas de células de estructura distinta, amplifica la luz antes de reflejarla.



COLIBRÍES

Los rayos de luz inciden sobre las plumas de los colibríes en una gran variedad de ángulos. La luz se refleja formando distintas coloraciones, que cambian dependiendo del movimiento de las alas.



torias, las ondas o bien se sumarán o se cancelarán mutuamente. De hecho, lo que sucede es que esa suma y cancelación se produce de forma periódica en la pantalla, dando lugar a lo que se conoce como «franjas de interferencia».

Pero con un dispositivo experimental como el anterior no siempre se obtiene un patrón de interferencia. Y es que, si la longitud de onda, que es la distancia entre dos crestas, es mucho menor que la distancia entre las rendijas, el efecto no se observa. Fue por esta razón por la que se tardó tanto en descubrir que la luz era una onda: su longitud de onda es de solo unos cientos de nanómetros, lo que hacía muy difícil observar patrón de interferencia alguno, llevando a la conclusión errónea de que debía de estar constituida por partículas.

Hemos visto dos versiones diferentes del experimento de la doble rendija. Por un lado, usando balines uno obtiene dos franjas, centradas en las trayectorias que unen la pistola con cada rendija. Usando ondas, se obtiene un patrón de interferencia como el de la figura 5.

Ahora nos lanzaremos a explicar la versión cuántica del experimento de la doble rendija. Puede hacerse usando desde fotones, que son las partículas de las que está hecha la luz, hasta moléculas. En este caso supondremos que usamos electrones, las partículas de carga negativa que orbitan los núcleos atómicos.

El dispositivo experimental para la doble rendija cuántica (figura 6) es el mismo que en el caso de los balines; la única diferencia es que la pistola es sustituida por un cañón de electrones y que la separación entre rendijas es mucho menor. Además, hay que cambiar la pantalla por algo que pueda detectar los impactos de los electrones, por ejemplo una pantalla de fósforo que emita un pulso de luz al ser golpeada por un electrón.

Como antes, el cañón de electrones dispara las partículas, que tras pasar por las rendijas chocan contra la pantalla, en la que aparece el punto que marca el impacto. Dado que los electrones son partículas, uno esperaría encontrar un patrón similar al de los balines: dos franjas concentradas en las trayectorias que unen el cañón de electrones con las rendijas. Sin embargo, y para la sorpresa general, se obtiene algo parecido al dibujo de la figura 7.

FIG. 6

Cañón de electrones

Doble rendija en la que se emplean electrones en lugar de balines. La pantalla de fósforo emite un pulso de luz cada vez que un electrón impacta contra ella.

Pantalla de fósforo

FIG. 7

Patrón de interferencia de electrones en la pantalla de fósforo. Curiosamente, la distribución de los impactos difiere mucho de la que se obtenía con los balines.

Como puede apreciarse, los impactos de los electrones crean lo que parece ser un patrón de interferencia como el de las ondas de agua. Sin embargo, los electrones en sí parecen ser partículas, porque al impactar contra la pantalla dejan solo una marca de impacto. Esto parece algo paradójico: si los electrones son partículas ¿cómo pueden dar lugar a un patrón de interferencia? Al fin y al cabo, la interferencia entre ondas se produce al interactuar dos ondas provenientes de ambas rendijas. Pero el electrón es una partícula, así que solo puede pasar por una.

La primera explicación que uno puede plantearse es que los electrones, en grupo, se comportan como una onda. Quizá hay algo en el patrón de su movimiento colectivo que hace que sus impactos se distribuyan como en un patrón de interferencia. Para comprobar si esta hipótesis es cierta, uno puede decidir enviar los electrones de uno en uno, en cuyo caso el patrón de interferencia debería desaparecer. Al intentarlo se obtiene una secuencia parecida a la que se muestra en la figura 8.

Este resultado parece atentar contra todo sentido común. El hecho de que aparezca un patrón de interferencia indica que hay algo que está pasando por ambas rendijas, porque si no, no habría lugar a interferencia alguna. Pero tenemos un solo electrón: ¿cómo puede una sola partícula pasar por dos sitios a la vez? Y, si lo hace, ¿por qué no deja entonces más de una marca en la pantalla?

Parece que necesitamos esclarecer si, por alguna razón desconocida, el electrón pasa por las dos rendijas a la vez. Para ello, bastará con poner un detector de electrones en cada rendija, lo que nos permitirá saber si el electrón de verdad se desdobra o si, por el contrario, solo atraviesa una de ellas.

El nuevo experimento parece restaurar el sentido común: los detectores revelan que el electrón pasa por la rendija superior o por la inferior, pero no por las dos a la vez. Sin embargo, hay otro efecto inesperado. Al poner detectores en las rendijas, el patrón de interferencia desaparece y el electrón pasa a comportarse como un balón. Es decir: el detector en las rendijas nos dice por dónde pasa el electrón, pero entonces ya no obtenemos el resultado anterior. En cambio, si quitamos los detectores ya no

EL EXPERIMENTO DE YOUNG «REVERSIONADO»

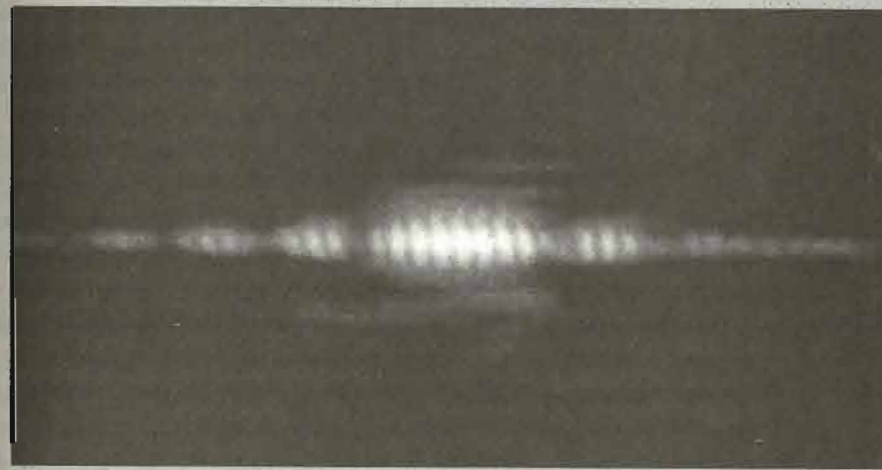
El experimento de la doble rendija es uno de los pocos experimentos clásicos que pueden realizarse con materiales disponibles en cualquier casa. Si en el experimento original de Young solo requería una vela y una hoja de papel, en este caso se utilizará un puntero láser que hará que el patrón de interferencia sea más fácil de obtener. Los punteros láser son ideales porque emiten luz de una sola longitud de onda, de forma que el patrón de interferencia se ve con más facilidad.

Dividiendo la luz

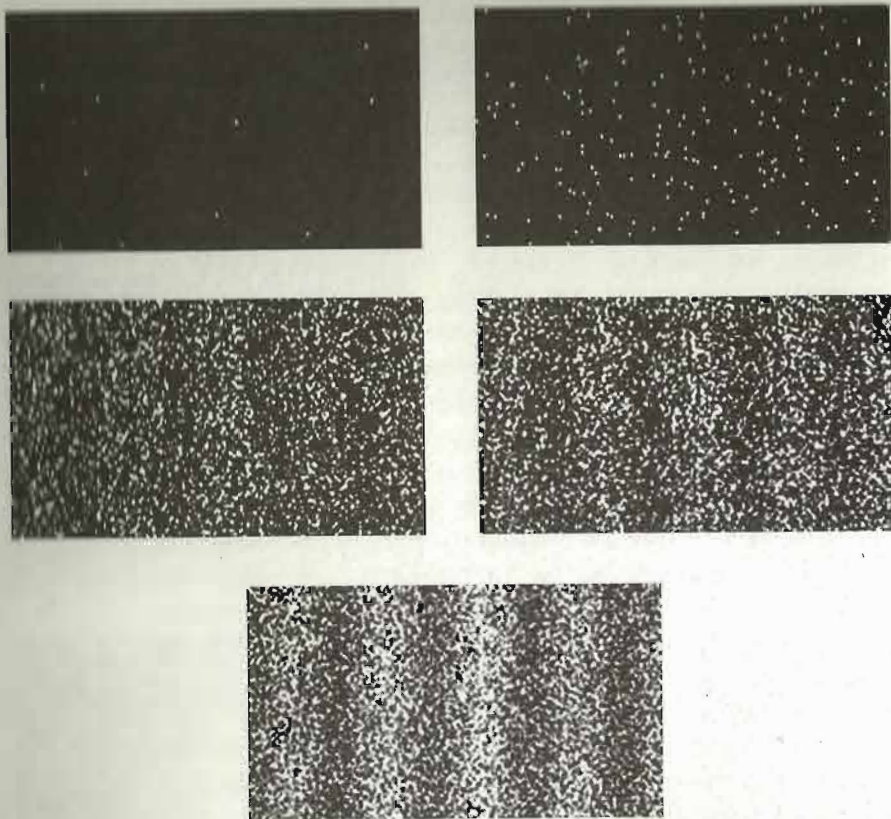
Procederemos a proyectar la luz de un puntero láser a través de un agujero (que habremos practicado con una chincheta o una aguja de coser) en una hoja de papel, preferiblemente de color negro. Esto nos proporcionará un haz de luz lo suficientemente pequeño. Como la longitud de onda de la luz roja es muy pequeña, en lugar de usar dos rendijas para separarla, usaremos un naipe, que tiene las dimensiones perfectas. Así pues, la luz se dividirá en dos haces, cada uno de los cuales pasará por un lado del naipe.

El fenómeno de la difracción

Para observar el patrón de interferencia solo se necesita una superficie blanca, como una pared. Si se ha preparado todo correctamente, tendría que observarse un patrón de interferencia como el que se muestra en la imagen: además del clásico patrón de interferencia, se observa también que los puntos brillantes se agrupan en conjuntos mayores. A este efecto se le llama *difracción*.



En este patrón de doble rendija se aprecian las franjas de interferencia en la parte superior del patrón de difracción. Los puntos se agrupan formando conjuntos.



Las imágenes muestran diferentes estados del experimento de la doble rendija enviando los electrones primero de uno en uno (fotografía superior) y luego de forma cada vez más numerosa. A medida que van chocando más electrones contra la pantalla, emerge un patrón de interferencia.

sabemos por qué rendija ha pasado el electrón, pero reaparece el patrón de interferencia.

Resumamos: cuando no hay detectores en las rendijas, el electrón parece comportarse como una onda, ya que crea un patrón de interferencia en la pantalla; cuando ponemos detectores en las rendijas, el electrón pasa a comportarse como una

partícula y el patrón de interferencia desaparece. Es decir: el electrón se comporta de forma distinta dependiendo de cómo se lo observe.

LA DOBLE RENDIJA Y EL PAPEL DEL OBSERVADOR

El resultado anterior resulta chocante incluso hoy en día, casi un siglo después de su descubrimiento. ¿Cómo puede ser que una observación modifique el comportamiento de una partícula? En la actualidad existe una variedad de explicaciones sobre qué es lo que realmente pasa en el experimento de la doble rendija. Se escoja la interpretación que se escoja, uno se ve obligado a replantearse por completo el papel del observador en la física.

Que el observador modifica lo que observa es algo que debería ser conocido por cualquier científico social. Por ejemplo, si queremos saber la opinión de la población sobre su presidente, podemos hacer una encuesta al respecto, pero existe la posibilidad de que la formulación de las preguntas altere la respuesta de la gente. Este fenómeno, común en disciplinas como la sociología, era completamente inesperado en física. Al fin y al cabo, una mesa que mide metro y medio tiene esa medida independientemente de quién y cómo la mida.

Una medida dependiente del observador es mucho más catastrófica en física que en las ciencias sociales. Cuando lo decimos desde una perspectiva psicológica no estamos afirmando que la gente piensa una cosa u otra dependiendo de cómo se le pregunte, sino solo que nuestro acceso a esos pensamientos puede tener un sesgo. Es decir: la realidad no depende del observador, solo nuestro acceso a ella. Sin embargo, al aceptar que un electrón es una onda o una partícula según cómo lo midamos, parece que estemos diciendo que la realidad misma depende de quién la mida. Esto es algo inaceptable para un físico, cuyo trabajo tiene como base la premisa de que existe una realidad cognoscible y que esta es la misma para todos.

El experimento de la doble rendija obligó a los físicos a incluir el propio acto de medir en las leyes de la física. Esto era algo

impensable en mecánica clásica: la segunda ley de Newton dice que la fuerza es el producto de la masa por la aceleración, no que lo es dependiendo de cómo y cuándo se midan las cantidades. En la mecánica cuántica, en cambio, el acto de medir es una parte fundamental de la teoría.

Si tenemos en cuenta que la observación forma parte del sistema físico estudiado, las dos posibilidades para el experimento de la doble rendija parecen ahora dos situaciones claramente distintas: por un lado, tenemos una situación donde la posición del electrón se mide cuando impacta en la pantalla; por otro, la posición se mide al pasar por la rendija. Dado que ambas situaciones son distintas, parece natural encontrar resultados distintos.

El esquema del experimento de la doble rendija se puede generalizar para otras situaciones. Para empezar, si no ponemos detectores en las rendijas el electrón parece comportarse como una onda. Esto nos da una idea: ¿y si el electrón realmente fuese una onda antes de chocar con la pantalla? Sin embargo, al chocar contra la pantalla, su estado cambia y lo observamos como si fuera una partícula.

Así pues, parece que necesitamos dos conjuntos de leyes: el primero para describir el comportamiento de la onda antes de chocar con la pantalla; el segundo para pasar del estado ondulatorio del electrón a lo que de hecho medimos.

Este planteamiento llevó a la comunidad física de entonces a la idea de que las partículas subatómicas son *ondas de probabilidad*. Dicho de otra forma, las partículas se comportan como ondas hasta que son medidas, momento en el que pasan a tener una posición determinada. Más tarde se verá que la posición es solo una de las muchas propiedades que pueden medirse y que, de la misma forma que un electrón puede tener varias posiciones a la vez en su estado ondulatorio, también puede tener varias energías o estados de rotación.

Una onda de probabilidad, pues, se comporta como una onda hasta que es medida. En ese instante la naturaleza escoge una de las diferentes posibilidades como resultado de la medida. La razón por la que se las llama ondas de probabilidad es que la «amplitud» de la onda, el equivalente a la altura del agua, nos

da información sobre la probabilidad de encontrar a la partícula en ese estado. Para ser exactos, la probabilidad de encontrar al electrón en una determinada posición viene dada por el cuadrado de la amplitud de la correspondiente onda de probabilidad.

A la ecuación que describe la onda asociada a una partícula antes de ser observada se la llama *función de onda* y su comportamiento viene dictado, en el caso no relativista, por la llamada *ecuación de Schrödinger*. Al proceso por el cual una medida hace que la partícula pase a tener solo un estado de los muchos posible se le llama el *colapso de la función de onda*, ya que parece que la onda se colapse y pase a ocupar solo un pequeño intervalo de posibilidades. Después de efectuar la medida, sin embargo, la partícula volverá a comportarse como una onda hasta que sea medida de nuevo.

Podemos resumir lo que hemos descubierto en los siguientes cinco puntos:

1. Antes de ser observada, una partícula se comporta como si fuera una onda.
2. El comportamiento de esa onda viene dado por la ecuación de Schrödinger.
3. Al ser observada, la función de onda *colapsa* y se mide solo una de las posibilidades, de forma aleatoria. Esto se suele conocer como el *postulado del colapso*.
4. La probabilidad de observar a la partícula en un cierto estado viene dada por el cuadrado de la amplitud de la parte de la función de onda asociada a esa posibilidad. Esto se conoce como la *regla de Born*.
5. Después de ser observada, la partícula sigue evolucionando como una onda hasta que es observada de nuevo.

Ahora que conocemos el efecto que una medida tiene sobre la evolución del electrón, podemos entender por qué colocar detec-

LA ECUACIÓN DE SCHRÖDINGER Y EL EFECTO TÚNEL

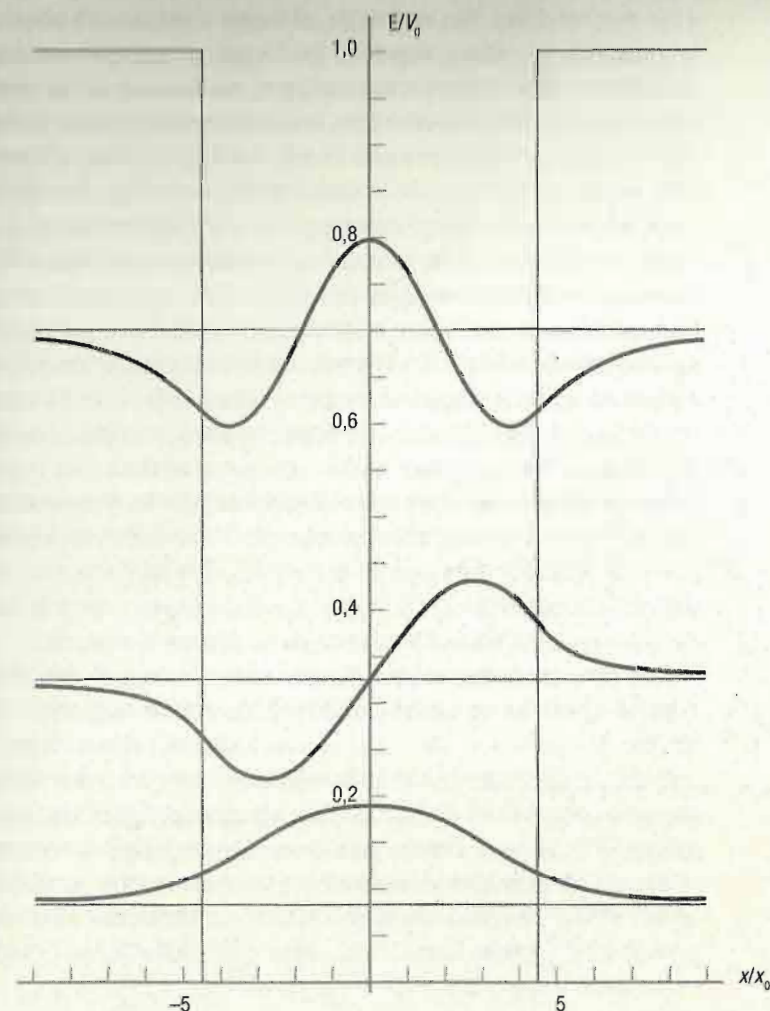
La evolución de la función de onda de una partícula viene dada por la ecuación de Schrödinger y fue publicada por primera vez en 1926. La célebre ecuación puede usarse tanto para predecir el comportamiento de un electrón en una doble rendija como para calcular la forma de las órbitas que describen los electrones en un átomo. También nos dice cómo evoluciona un sistema cuántico con el tiempo. Sin embargo, en la mayoría de sistemas cuánticos que se estudian, los físicos están más interesados en conocer los diferentes niveles de energía posibles que en la evolución temporal. Para ello se suele usar una versión especial de la ecuación de Schrödinger llamada *ecuación de Schrödinger independiente del tiempo*. Se trata de una *ecuación diferencial*: una ecuación donde lo que se desea conocer es una función con ciertas características y lo que se nos da es cómo cambia, ya sea en el espacio o en el tiempo. Suele escribirse como:

$$\hat{H}\Psi = E\Psi.$$

Esta formulación simple esconde una enorme complejidad. El símbolo H se llama *hamiltoniano* y es un «operador»: una suerte de máquina a la que, tras introducirle una función, escupe otra. Su significado es el siguiente: si introducimos la solución de nuestra ecuación en el operador hamiltoniano, nos devolverá exactamente la misma función, multiplicada por un número que resulta ser la energía, E . No hay duda de que la ecuación de Schrödinger puede complicarse mucho, en función del objeto de estudio. Los sistemas más estudiados son el llamado «oscilador armónico», que equivale a una masa sujeta a un muelle, y el átomo de hidrógeno. Hay un caso donde la resolución de la ecuación de Schrödinger es relativamente sencilla: se trata del denominado «pozo de potencial», que representa una situación donde la partícula está atrapada en una cierta región del espacio. En la figura se muestran tres posibles soluciones.

Curiosos movimientos particulares

La probabilidad de encontrar a la partícula en un punto determinado viene dada por el cuadrado de la altura de la curva. Lo que resulta sorprendente es que, en este caso, la probabilidad de encontrar a la partícula más allá de la barrera no es cero: las partículas cuánticas son capaces de traspasar objetos sólidos. A esto se le llama «efecto túnel», algo que, lejos de ser esotérico, sucede constantemente. Se trata de un fenómeno cuántico a consecuencia del cual una partícula, violando los principios de la mecánica clásica, penetra en un área que, en teoría, no debería poder atravesar al no tener la energía cinética suficiente; pero, sin embargo, lo consigue. Por ejemplo, los electrones de un átomo son capaces de atravesar el núcleo sin pasar por él: es decir, pueden desaparecer en el lado derecho y aparecer en el izquierdo. Si eso no sucediera, los elementos químicos tendrían propiedades muy distintas a las que conocemos.



Soluciones a la ecuación de Schrödinger para el pozo de potencial. La altura de la curva al cuadrado nos da la probabilidad de encontrar a la partícula en esa región.

tores en las rendijas altera el patrón de interferencia. Pongamos que un electrón sale del cañón. Al principio, el electrón se comportará como una onda normal y corriente y se dirigirá por igual a las dos rendijas. Sin embargo, al llegar a estas será observado y su función de onda colapsará. Dado que el colapso es aleatorio, será detectado o bien en la rendija superior o en la inferior, pero nunca en las dos a la vez. Una vez observado, pasará por la rendija y seguirá evolucionando como onda, pero esta vez no habrá otra onda proveniente de la otra rendija con la que interferir. Así pues, el patrón de interferencia ya no será observado.

Esta nueva forma de pensar da cuenta del resultado del experimento de la doble rendija, pero ha habido que pagar un precio. Por un lado, la mecánica cuántica introduce por primera vez la aleatoriedad en física. Se trata de una aleatoriedad muy diferente a la de un dado: un dado es, en principio, predecible, aunque en la práctica no lo sea. Por el contrario, la aleatoriedad en mecánica cuántica es fundamental: cada vez que se realiza una medida, el universo parece escoger un resultado aleatorio. No es un problema de nuestra ignorancia ni nada que se arregle con un ordenador más potente: las leyes de la física tienen un componente aleatorio. Esta conclusión fue difícil de digerir para un gran número de físicos, entre ellos Einstein, que murió sin aceptarla.

Por otro lado, hay un problema de índole quizá más filosófica aunque igualmente apremiante: ¿cómo puede depender la evolución del universo de un observador? En primer lugar, en el universo no siempre ha habido observadores conscientes. En segundo lugar, ¿cómo definimos un observador? ¿Necesitamos una persona? ¿O basta con un gato o una lagartija? ¿Y un ordenador? ¿Cuándo se produce exactamente el colapso de la función de onda? ¿Tiene sentido decir que el acto de observar algo modifica la realidad? ¿Existe la realidad antes y después de ser observada?

LA DECOHERENCIA CUÁNTICA: EL OBSERVADOR NO IMPORTA

La mecánica cuántica se basa en el concepto de observación, pero su definición no está de momento muy clara. ¿Cuándo se

produce una observación? La respuesta más obvia es que esta sucede cuando una persona es consciente del resultado de un experimento, es decir, en el momento de la percepción por parte del individuo. Por ejemplo, en el caso del experimento de la doble rendija la observación se produce cuando el experimentador ve el patrón en la pantalla, pero no antes.

Esta primera aproximación tiene varios problemas que se pueden ilustrar con uno de los experimentos mentales más famosos de la mecánica cuántica, el del llamado «gato de Schrödinger». Se trata de lo siguiente: supongamos que tenemos un gato metido en una caja con una pistola infinitamente silenciosa apuntándole a la cabeza. La pistola se disparará según un proceso radiactivo que tiene un cincuenta por ciento de posibilidades de suceder. Antes de abrir la caja ¿el gato está vivo o muerto?

Si aceptamos que la observación solo se produce cuando un observador humano la hace, la respuesta tiene que ser por fuerza que, antes de abrir la caja, el gato está un cincuenta por ciento vivo y un cincuenta por ciento muerto. Esto es debido a que la materia se comporta según el modelo ondulatorio antes de ser observada: de la misma forma que el electrón, como onda, pasa por ambas rendijas si no es observado, el gato está vivo y muerto a la vez si no se abre la caja.

Hay diferentes formas de solucionar la aparente paradoja del gato de Schrödinger. La más directa es simplemente aceptar que el gato está, en efecto, vivo y muerto a la vez y que creer lo contrario significa seguir anclado en formas desfasadas de pensar. Otra forma de mitigar el problema, aunque no de resolverlo, es a través de la llamada *decoherencia cuántica*. La idea de la decoherencia es que el acto de observar un sistema cuántico no está relacionado con la presencia de un observador consciente, sino con la interacción de este sistema con un cuerpo macroscópico, como un físico experimental, un gato o un detector de radiación. Si bien esto no resuelve todos los problemas filosóficos de la mecánica cuántica, sí arroja luz sobre algunos.

Nadie es solo él mismo.

ERWIN SCHRÖDINGER

Para entender la discusión que sigue, será necesario introducir la noción de *entrelazamiento cuántico*. En mecánica cuántica, se dice que dos partículas están entrelazadas si las propiedades que se miden en una dependen de las de la otra. La separación de la teoría en los procesos de evolución temporal y observación hace que se den fenómenos que pueden parecer poco intuitivos.

El sistema entrelazado más sencillo es el de dos partículas con *espín* opuesto. El espín puede entenderse como la cantidad de rotación de una partícula sobre sí misma. El espín suele representarse como una flecha que apunta arriba o abajo, según el sentido de la rotación (figura 9).

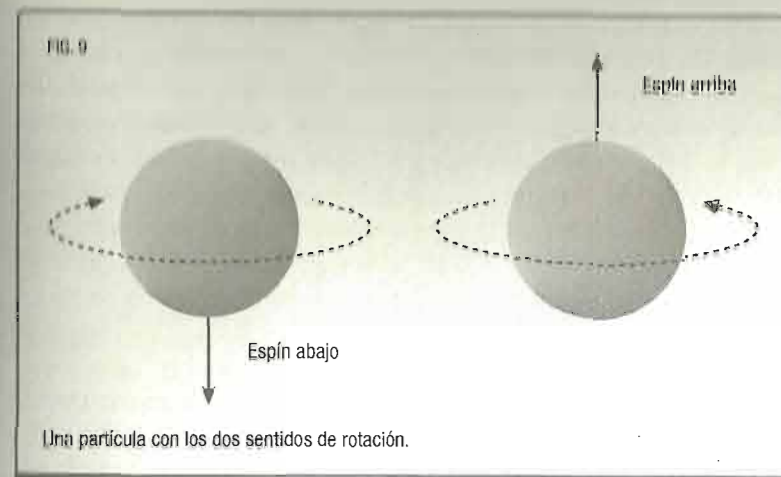
En un experimento es posible medir el espín de una partícula en una determinada dirección. Para un electrón se pueden obtener dos valores distintos, a los que se suele llamar informalmente como espín «arriba» o espín «abajo». En el caso de tener dos electrones, si ambos apuntan en la misma dirección, se suman; si lo hacen en direcciones opuestas, el espín suma cero.

Imaginemos que tenemos un sistema de dos partículas y sabemos que su espín total en una determinada dirección es cero. Bautizaremos a las dos partículas como A y B. En este caso, sabemos que, si el espín de la partícula A es hacia arriba, el de la partícula B tiene que ser hacia abajo y viceversa. Y no solo eso: dado que se trata de un sistema cuántico, antes de hacer una medida nuestro sistema se encontrará en los dos estados a la vez. Podemos representar esto de forma esquemática de la siguiente forma:

$$|\text{Estado del sistema}\rangle = |A \text{ arriba}\rangle |B \text{ abajo}\rangle + |A \text{ abajo}\rangle |B \text{ arriba}\rangle.$$

Esta notación algo informal para representar el estado del sistema es muy parecida a la que se usa para hacer cálculos en mecánica cuántica. En este caso, la suma indica que nuestra partícula está en los dos estados a la vez.

Si midiésemos el espín del sistema de partículas que hemos llamado A y B, la función de onda colapsaría y nuestro sistema pasaría a ser representado solo por una de las opciones,



dependiendo de qué obtuviéramos al medir. Una de las posibilidades es:

$$|\text{Estado del sistema después de la medida}\rangle = |A \text{ arriba}\rangle |B \text{ abajo}\rangle.$$

Se puede ver cómo el efecto de una observación es la eliminación de todos los términos de la suma excepto el que se ha medido. Este es el aspecto matemático del colapso de la función de onda. Hasta ahí, todo parece perfectamente normal. Sin embargo, podríamos tomar una de nuestras partículas y llevárnosla a la Luna, dejando la otra en el laboratorio. Ya que aún no hemos realizado medida alguna, el estado de nuestro sistema sigue dado por:

$$|\text{Estado del sistema}\rangle = |A \text{ arriba}\rangle |B \text{ abajo}\rangle + |A \text{ abajo}\rangle |B \text{ arriba}\rangle.$$

Supongamos ahora que el físico experimental del laboratorio decide medir el espín de la partícula A, que se ha quedado en la

ESTADOS, OPERADORES Y NORMALIZACIÓN

En la notación actual, el estado de una partícula se representa mediante una notación abstracta que popularizó el físico británico Paul Dirac (1902-1984). En esta notación, el estado de un sistema se representa mediante el símbolo denominado ket y que se suele escribir del siguiente modo:

$$|\Psi\rangle$$

Un ket es parecido a los vectores que se enseñan en la escuela secundaria. Así como un vector tiene tres componentes y se puede escribir como tres números separados por comas, un ket tiene tantas componentes como estados posibles de la partícula, normalmente infinitos. Sin embargo, a todos los efectos podemos leer el ket como algo que simplemente denota el estado del sistema. Supongamos, por ejemplo, que tenemos una partícula que puede encontrarse en una de dos energías. En ese caso, el estado del sistema será una suma de los dos posibles estados, es decir:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |\text{Energía 1}\rangle + \frac{1}{\sqrt{2}} |\text{Energía 2}\rangle$$

Puede verse que hemos añadido un factor de $1/\sqrt{2}$ a cada estado. Esto se debe a la regla de Born, que dice que la probabilidad de un estado viene dada por el número que lo multiplica al cuadrado. De este modo, tanto la probabilidad de la primera energía como la de la segunda será:

$$\text{Probabilidad} = \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} = 0,5.$$

Es decir, un 50 %. Como se ve, ambas probabilidades tienen que sumar uno: esto es lógico, ya que la probabilidad de que la partícula esté en un estado u otro tiene que ser exactamente del 100 %. Por supuesto, también puede suceder que un estado sea más probable que el otro. Por ejemplo, en el siguiente estado la segunda energía es más probable que la primera:

$$|\Psi\rangle = \frac{3}{5} |\text{Energía 1}\rangle + \frac{4}{5} |\text{Energía 2}\rangle$$

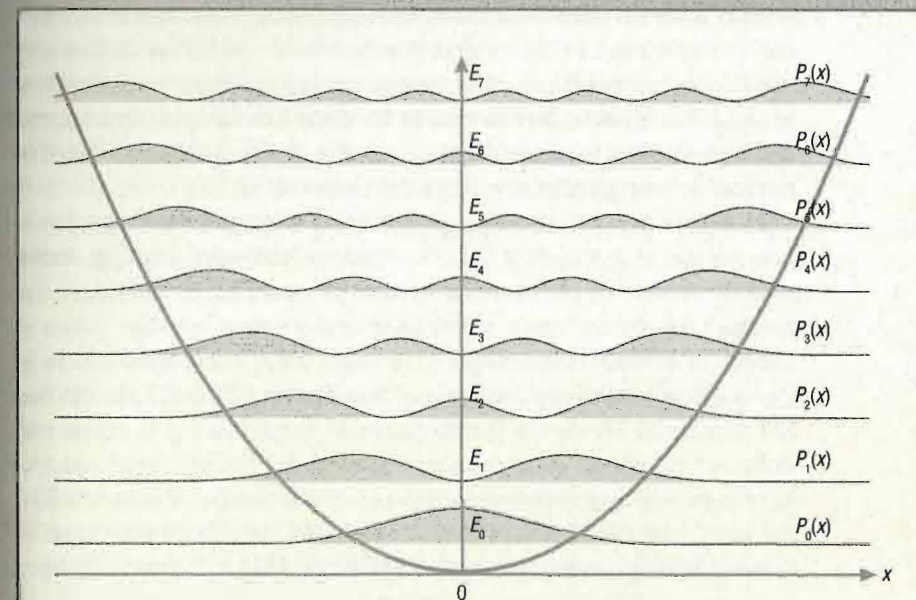
En este caso, la probabilidad de obtener la primera energía es:

$$\text{Probabilidad} = \left(\frac{3}{5}\right)^2 = \frac{9}{25} = 0,36.$$

○ un 36 %. Lo que significa que la de la segunda energía tendría que ser un 64 %:

$$\text{Probabilidad} = \left(\frac{4}{5}\right)^2 = \frac{16}{25} = 0,64.$$

Comprobémoslo en la gráfica de niveles de la figura.



Niveles de energía de un sistema cuántico. En general, una partícula se encontrará en una superposición de varios niveles posibles, hasta que su energía sea medida, en cuyo caso se asentará en el nivel energético que se haya medido.

Tierra, y obtiene «abajo». Inmediatamente, la función de onda colapsa y el estado del sistema viene dado por:

$$|\text{Estado del sistema}\rangle = |\text{A abajo}\rangle |\text{B arriba}\rangle.$$

Como puede verse, al medir el espín de la partícula A, el de la partícula B ha quedado inmediatamente determinado. ¡Pero la partícula B está en la Luna! Es decir: una medición realizada en la Tierra tiene un efecto inmediato en cualquier partícula entrelazada con la que estemos midiendo, esté donde esté. Por eso se suele decir que la mecánica cuántica es una teoría no local, en el sentido de que hay efectos que parecen propagarse instantáneamente. Esta predicción aparentemente absurda se ha visto verificada numerosas veces en el laboratorio y se ha utilizado incluso para transmitir información, aunque a velocidades menores que la de la luz. A este fenómeno se le suele llamar «teletransporte cuántico» aunque en realidad lo que se teletransporta es una propiedad de una partícula y no la partícula en sí.

El entrelazamiento cuántico es esencial para resolver o al menos mitigar el problema del observador. Para ello, hay que darse cuenta de que el observador mismo es un sistema cuántico: no estamos «autorizados» a considerar el sistema a estudiar como si estuviese aislado del mundo. Esto es un salto cualitativo desde la física clásica: ya no es posible separar mente y materia al estudiar la naturaleza. La mente forma parte de la naturaleza y, como tal, debe ser tenida en consideración al modelar los procesos naturales, ya que interactúa con esos procesos por medio de la medición. Se ha cerrado el círculo: de un mundo que se comporta como la mente, se ha pasado a una mente que sigue las leyes de la física y que es una parte inseparable de nuestro modelo explicativo.

Esta forma de analizar el problema logra arrojar cierta luz sobre lo que realmente sucede durante el experimento de la doble rendija. Centrémonos primero en la situación en la que observamos al electrón en la rendija misma, antes de impactar contra la pantalla. Nos interesa saber qué tipo de procesos físicos suceden

en el acto de hacer una medida: es decir, no basta afirmar «aquí tenemos un aparato que detecta si pasa un electrón», sino que nos vemos obligados a analizar el aparato mismo como parte del proceso que estamos intentando entender.

Detectar un solo electrón no es un proceso sencillo. Una corriente típica, de unas décimas de Ampère, contiene del orden de 10^{18} electrones, es decir, un uno seguido de dieciocho ceros o un trillón. Para detectar un solo electrón debemos ser capaces de amplificar su señal drásticamente, convirtiendo esa pequeña corriente en otra que será millones de veces mayor y detectable por nuestros aparatos.

Así, cuando un electrón es detectado lo que realmente está pasando es que el estado del electrón queda entrelazado con el estado de los miles de millones de partículas que forman el detector. Esto no significa que el electrón interactúe con el detector dependiendo del estado en que esté: el electrón, como onda de probabilidad, interactúa con el detector, también como onda de probabilidad, de forma que algunos estados posibles del detector quedan ligados a otros estados posibles del electrón, pero aún se tiene un estado tipo onda donde todos los estados posibles se dan. Esto se puede describir esquemáticamente de la siguiente manera:

$$\begin{aligned} |\text{Estado de todo el sistema}\rangle &= |\text{Electrón arriba}\rangle \\ &|\text{Detector arriba}\rangle + |\text{Electrón abajo}\rangle |\text{Detector abajo}\rangle. \end{aligned}$$

Como puede verse, tenemos otra situación de entrelazamiento, en este caso entre el electrón y el aparato de medida. En otras palabras: medida y entrelazamiento son la misma cosa. Realizar una medida cuántica equivale a entrelazar el estado del electrón con el estado del aparato de medida. Sin embargo, el proceso no acaba aquí: justo después de la medida, el entorno en el que nos encontramos también queda entrelazado con el electrón original. Por ejemplo, si el detector dice «arriba» saldrán de este una multitud de fotones que se esparcirán por la habitación en cues-

tión de milésimas de segundo. Esos fotones colisionarán con átomos y estos a su vez interactuarán con otros, creando una cascada que va mucho más allá del proceso original. Podemos modelar el sistema de la siguiente forma:

$$\begin{aligned} |\text{Estado de todo el sistema}\rangle &= |\text{Electrón arriba}\rangle \\ &|\text{Detector arriba}\rangle |\text{Entorno arriba}\rangle + |\text{Electrón abajo}\rangle \\ &|\text{Detector abajo}\rangle |\text{Entorno abajo}\rangle \end{aligned}$$

En estos momentos, parece que nos encontremos en exactamente la misma situación que al principio: tenemos dos estados posibles así que ¿por qué no podemos observarlos a la vez? Si el electrón se comporta de esa forma, también tendría que poder hacerlo nuestro entorno. Aquí entra en juego de nuevo el concepto de la decoherencia cuántica, basada en el concepto clásico de decoherencia entre dos ondas.

Recordemos el experimento de la doble rendija con ondas: en ese caso, las dos ondas interferían en la pantalla, creando un patrón de interferencia. Sin embargo, eso no siempre pasa: si la separación entre las rendijas es mucho mayor que la longitud de onda, no se observa interferencia alguna. De hecho, las condiciones para la interferencia son inmensamente restrictivas: si las dos ondas que salen de cada rendija no se parecen lo suficiente, tampoco se observará patrón de interferencia alguno. Eso sucede porque, aunque en un momento dado puede darse interferencia tanto constructiva como destructiva, esto cambia tan rápido que, de media, no se observa ningún efecto, al contrario que cuando ambas ondas tienen la misma longitud de onda y provienen de la misma fuente. Cuando las ondas no son lo suficientemente parecidas se las llama «decoherentes». Una condición necesaria para observar interferencia es que las ondas sean muy parecidas o, en términos físicos, «coherentes».

Si ahora volvemos al ejemplo del experimento de la doble rendija, veremos que nuestro sistema al completo, incluyendo electrón, aparato de medida y entorno, se ha dividido en dos partes distintas: en una, el electrón ha pasado por la rendija superior y el aparato de medida así lo indica; en la otra, el electrón ha pasado

por la otra rendija y, de nuevo, el aparato de medida se hace eco de ello. Pero, además de eso, en ambos casos ha habido una serie de cambios en nuestro entorno a los que hemos perdido la pista completamente: no tenemos ni idea de en qué se diferencian los dos laboratorios ni tenemos forma alguna de medirlo. La información, a efectos prácticos, se ha perdido y no podemos recuperarla.

Nos vemos obligados a ignorar el entorno y fijarnos solo en lo que podemos medir, que es lo que nos dice el detector. A todos los efectos, nuestra función de onda sigue teniendo dos partes: electrón arriba y electrón abajo. Pero ambas están entrelazadas con dos estados de nuestro entorno que no podemos medir. No solo eso: debido a que la cantidad de cambios es enorme, en la práctica las dos partes de la función de onda se han transformado en decoherentes: ya no pueden afectarse mutuamente. La interferencia ha desaparecido. Esa decoherencia cuántica se produce de forma muy rápida al interactuar un sistema microscópico con otro macroscópico: por ejemplo, en el caso del gato de Schrödinger pasará casi de inmediato. Si en nuestro experimento de la doble rendija solo observamos al electrón al chocar con la pantalla, las dos funciones de onda se parecerán lo suficiente como para interferir la una con la otra, creando un patrón de interferencia. En cambio, al observar al electrón en la rendija, la coherencia queda dañada y el patrón de interferencia desaparece. Como corolario, vemos que una medida no requiere de un observador consciente: todo lo que necesitamos es una interacción con un sistema lo suficientemente grande.

Sin embargo, el problema de la medida sigue sin solucionarse. La decoherencia explica por qué los estados disponibles se limitan a los que puede medir nuestro detector, pero no nos dice el motivo por el cual colapsa la función de onda. ¿Por qué no obtenemos ambos resultados al realizar una medida? ¿Por qué obtenemos uno y no el otro? ¿Cómo escoge la naturaleza entre ambos? La solución a estas cuestiones lleva decenios debatiéndose. Para resolverlas hay que añadir algo a la decoherencia: una interpretación de lo que significa. Hay dos principales contendientes con una visión del universo radicalmente distinta que, sin embargo, hacen las mismas predicciones experimentales.

Uno de los problemas más publicitados de la gravitación de Newton fue que contenía lo que se ha dado en llamar «acción a distancia»: la habilidad de los planetas de comunicar una fuerza a través de distancias enormes sin ningún mecanismo aparente. Newton salió al paso de las críticas con su famosa frase «no invento hipótesis», una cita que se usa con frecuencia para ilustrar el papel de la ciencia como un constructo cuyo cometido no es dar explicaciones, sino predecir correctamente los resultados experimentales. Una teoría puede ser atacada si no predice los fenómenos correctos, pero no si contiene un supuesto filosóficamente cuestionable. La ciencia es ciencia, no filosofía.

Por supuesto, la realidad es siempre un poco más prosaica. Borradores del manuscrito de Newton revelan que el físico barajó varias hipótesis sobre las causas de la gravitación universal antes de decantarse por no hacer ninguna. Está en el ADN de los científicos buscar explicaciones a los fenómenos que observan, aunque a veces no quede más remedio que aceptar que no tienen ninguna.

En mecánica cuántica pasa algo parecido, pero las posiciones son, si cabe, más enconadas. En el caso de Newton, no resulta tan difícil aceptar la existencia de una fuerza que se transmite a distancia; es mucho más difícil asentir y mirar para otro lado con la mecánica cuántica, donde el conflicto con el sentido común es mucho más obvio.

A lo largo del siglo xx aparecieron diferentes explicaciones más o menos filosóficas para los resultados de la mecánica cuántica. A esas explicaciones se las suele llamar «interpretaciones» de la mecánica cuántica, para distinguirlas de la teoría en sí y de sus predicciones, con las que todo el mundo está de acuerdo. Las interpretaciones se usan como una herramienta explicativa, pero el aparato matemático es el mismo en todos los casos. Hay que recalcar también que algunos físicos llegan a tachar a los proponentes de la interpretación contraria como acientíficos y embaucadores. En lo que sigue se tratará de exponer las dos interpretaciones más populares sin entrar en valoraciones.

Como se ha visto, tenemos una teoría con dos fases: en la primera, la materia se comporta como una onda, en cuya amplitud está codificada la probabilidad de encontrar a la partícula o sistema en un cierto estado. En la segunda se realiza una medida y la onda se colapsa, dando lugar a la medición realizada. A pesar de que es posible predecir la probabilidad de un determinado resultado, es imposible predecir de antemano qué resultado va a obtenerse. Para realizar una medida es necesario que un observador la lleve a cabo, lo que crea la incómoda necesidad de definir a qué nos referimos cuando decimos «observador».

Definir su papel fundamental dentro de la teoría llevó a la comunidad de físicos a plantear la primera y más popular interpretación de la misma, llamada *interpretación de Copenhague* al ser la favorita del físico Niels Bohr (1885-1962), natural de la capital de Dinamarca. Esta interpretación es antirrealista, en el sentido de que niega la realidad de la función de onda y, en cambio, admite la del observador. Según la interpretación de Copenhague, la función de onda no es nada más que un artificio matemático que nos da información sobre lo único real a lo que tenemos acceso, que es el resultado de una medida. La función de onda representa la información subjetiva que el observador tiene del sistema, nada más. Por eso, después de realizar una medida, debe ser actualizada con la nueva información. Discutir sobre si lo que pasa entre medida y medida es real, no solo no es productivo, sino que además es acientífico.

La interpretación de Copenhague es antigua y ha sido completada, aunque no oficialmente, con otros avances como la decoherencia. Además, los proponentes de Copenhague no son un grupo homogéneo sino que adoptan una posición más o menos realista según sus propias convicciones.

Una versión extrema de Copenhague que sirve para ilustrar la idea básica es la famosa cita de Richard Feynman (1918-1988), «calla y calcula». Según Feynman, si los físicos dejaran de perder el tiempo divagando sobre la realidad subyacente y se dedicasen a desarrollar el aparato matemático de la teoría, la física avanzaría mucho más rápido. Al fin y al cabo, las teorías físicas son solo modelos matemáticos que dan cuenta de resultados ex-

perimentales, que no tienen por qué representar lo que de verdad sucede. El universo podría estar hecho de queso verde y, mientras su comportamiento fuera dado por las leyes de la mecánica cuántica, eso no tendría importancia alguna. La física newtoniana afirmaba que existía una fuerza invisible entre planetas; en la gravedad de Einstein, esa fuerza se debe a una distorsión del espacio, mientras que en la gravedad cuántica es transmitida por unas partículas llamadas gravitones. ¿Quién tiene razón? Esta es una pregunta que no podemos responder: lo que sí podemos argumentar es qué teoría hace las predicciones más exactas.

Una característica curiosa de la interpretación de Copenhague es que sí asume la realidad de algo: la del observador consciente que lleva a cabo la medida y la del resultado de la medida en sí. En este sentido, se considera que la versión de Copenhague es «subjetiva» ya que renuncia a una realidad objetiva y externa al observador, dada por la función de onda, y la sustituye por la realidad de cada observador al realizar una medida. Esto resulta difícil de aceptar para algunos físicos, que creen preferible abandonar otros supuestos antes que la realidad unívoca del universo que los contiene. Algunas de las cuestiones que estos físicos podrían plantearse y que otros no considerarían lícitas son: ¿cómo podemos explicar la consciencia a partir de la física si asumimos que la física se refiere a la consciencia? ¿Qué estatus tiene la función de onda? Si no es real ¿qué representa? ¿De verdad tiene sentido afirmar que no existe la realidad antes de ser observada? ¿Qué pasa si hay múltiples observadores observando fenómenos distintos?

La teoría de múltiples universos de Hugh Everett

Hugh Everett III (1930-1982) intentó responder a estas preguntas con lo que hoy se ha dado en llamar la «interpretación de múltiples universos». Se trata de una interpretación que cuenta con proponentes entusiastas, como el físico británico de origen israelí David Deutsch, investigador de la Universidad de Oxford, Reino Unido, o el estadounidense Sean Carroll, cosmólogo en el

Instituto de Tecnología de California, el Caltech, pero también con detractores virulentos que la consideran acientífica, como el físico teórico Peter Woit, de la Universidad de Columbia, en Estados Unidos. La interpretación de múltiples universos puede verse como el reverso de Copenhague: mientras que una cree que la función de onda no es real pero el observador sí lo es, en múltiples universos es la función de onda la que es real, mientras que el observador solo es consciente de una parte muy limitada de lo que pasa en el universo.

Supongamos que tenemos un sistema cuántico, por ejemplo un electrón con su espín en una determinada dirección, cuyos únicos valores posibles son «arriba» y «abajo». Al principio no sabemos qué espín tiene el electrón, así que su función de onda será una suma de «arriba» y «abajo». Según Copenhague, esa función de onda representa solo nuestra información sobre el sistema: no sabemos si se encuentra en un estado o en el otro, así que debemos considerarlos los dos. Según la interpretación de múltiples universos, en cambio, la función de onda es una representación fidedigna del estado del electrón.

Después de un tiempo, decidimos realizar la medida. Como se ha visto, nuestro aparato de medida y nuestro entorno quedarán entrelazados con el electrón, de forma que tendremos una función de onda que incluya a ambos y que tendrá dos posibilidades: la detección del electrón hacia arriba y la detección del electrón hacia abajo.

Según Copenhague, la función de onda nos dice simplemente que hay dos posibles resultados de la medida: es decir, nos da información sobre lo que vamos a ver, pero no determina qué pasará. El resultado de la medida será aleatorio y vendrá dado por la regla de Born. Preguntarse sobre lo que ha pasado antes de la medida o lo que pasará después no tiene sentido.

La interpretación de múltiples universos niega el colapso de la función de onda. La función de onda es real y se rige por la ecuación de Schrödinger, haya un observador o doscientos. Así

La filosofía es solo una forma de hablar de los descubrimientos que ya se han realizado.

PAUL DIRAC

que sigue teniendo dos partes: una en la que se ha detectado el electrón hacia arriba y otra en la que se ha detectado hacia abajo. El laboratorio, como el electrón, se desdoblará en dos posibilidades: en una, aparecerá el resultado «arriba». En la otra, aparecerá el resultado «abajo».

Sin embargo, esta explicación parece estar en conflicto con la observación. Si observamos ambos casos ¿por qué somos solo conscientes de uno? La respuesta está en la decoherencia cuántica: ambas partes de la función de onda son decoherentes y no pueden interactuar entre sí, de forma que son invisibles la una a la otra. Parece que el universo se haya dividido en dos, aunque de hecho solo hay una función de onda. Decimos que el universo se ha desdoblado en dos «universos paralelos» en el sentido de que parece haber dos universos que no interactúan entre sí.

La interpretación de múltiples universos es objeto de críticas frecuentes por parte de algunos integrantes de la comunidad física. En lo que sigue se resumirán algunas de las objeciones a la interpretación y, si cabe, la respuesta de los proponentes de la teoría.

La acusación más común a la interpretación de múltiples universos es que es acientífica porque no hace predicciones contrastables. Sin embargo, esto es una verdad a medias: la interpretación de múltiples universos es una versión de la mecánica cuántica y, como tal, está sujeta a los mismos exámenes de verificabilidad que esta. Sí es verdad que es imposible distinguir experimentalmente entre las interpretaciones de Copenhague y la de los múltiples universos, pero la misma crítica podría hacerse a cualquiera de las interpretaciones. Si una teoría tiene varias interpretaciones compatibles con ella, o bien son todas acientíficas o ninguna lo es. Lo que es cierto es que la interpretación de Copenhague es la más minimalista en sus afirmaciones: es decir, Copenhague es lo más cercano que hay a no hacer interpretación alguna y, en ese sentido, puede considerarse más científica. Dicho de otra forma: la interpretación de Copenhague hace el mínimo de supuestos sobre la realidad para poder dedicarse a hacer ciencia.

Por otro lado, se ha criticado a la interpretación de múltiples universos por no respetar la navaja de Ockham, un principio me-

todológico atribuido al fraile franciscano Guillermo de Ockham (1280-1349) que afirma que «en igualdad de condiciones, la explicación más sencilla suele ser la más probable». Es decir: no deben multiplicarse las entidades de una explicación sin necesidad. En este sentido, parece claramente anti-ockhamiana: postula una infinidad de universos no observables que se reproducen cada vez que se hace una medida. Esto es debatible: la aparición de los múltiples universos es una consecuencia del rechazo al colapso de la función de onda y no son introducidos *ad hoc*. De hecho, la frase «múltiples universos» lleva a confusión: lo que hay es una única función de onda que se divide en una serie de partes no interactuantes.

Existen, sin embargo, críticas de peso a la interpretación de múltiples universos. La más devastadora es que no puede explicar por qué las partes de la función de onda nos dan las probabilidades de obtener cada medida. A pesar de que ha habido intentos para corregir este punto, por ejemplo por parte de Deutsch, su éxito hasta la fecha es puesto en entredicho por una parte sustancial de la comunidad física.

MENTES CLÁSICAS ASEDIADAS POR FENÓMENOS CUÁNTICOS

La visión de la consciencia que tienen ambas interpretaciones es diametralmente opuesta, a pesar de que hacen las mismas predicciones. En Copenhague, el acto de observación es real y la existencia de una realidad externa al observador, al menos de una realidad determinada, es una pregunta que no estamos autorizados a hacernos. De forma algo metafórica podríamos decir que la realidad es borrosa mientras que los observadores están bien definidos. Por otro lado, en la interpretación de múltiples universos la función de onda es real, mientras que el colapso es solo aparente.

Tanto los múltiples universos como la percepción subjetiva planteada en Copenhague tienen implicaciones filosóficas sobre la consciencia que nos alejan de la física clásica del siglo XIX. La visión de Copenhague pone el foco en lo que podemos saber

en lugar de en qué pasa realmente. Un filósofo diría que es más epistemológica que ontológica; nos interesa qué podemos saber y nos da igual lo que hay.

La ciencia, en esta forma de ver, no se ocupa del universo sino de la mente. Su cometido no es describir el mundo, sino predecir las percepciones del individuo. Algunas teorías se ayudarán para ello de una realidad compartida por todos; otras no necesitarán tal constructo o lo harán en menor medida. Pero da lo mismo: mientras nuestra teoría prediga las observaciones correctas, no podrá hacersele objeción alguna. Desde esta lente, parece natural referirse a la función de onda como un modelo matemático de la información que posee un observador. Dado lo que sabemos en cierto momento ¿podemos predecir lo que sabremos luego?

Un realista se quejaría, diciendo que el sujeto tiene que existir en algún lugar. Un antirrealista no lo negaría: simplemente afirmaría que la cuestión es baladí. De interés filosófico, tal vez, pero no científico. La creencia de que el sujeto tiene que existir en un universo determinado que lo contiene es una cuestión personal de preferencias o, si se quiere, de prejuicios sobre cómo tiene que ser el mundo. Pero la ciencia no se ocupa de estos últimos.

La interpretación de múltiples universos tiene implicaciones igualmente chocantes para la consciencia, aunque por motivos muy distintos. Por ejemplo, predice la existencia de multitud de copias de cada persona, una para cada observación posible. Esto crea un problema: si existen tantas copias ¿por qué solo somos conscientes de ser una de ellas? Si una de nuestras copias ha ganado la lotería, ¿por qué no somos conscientes de ello?

Una posible respuesta es que todas nuestras copias son conscientes de existir. Lo único que pasa es que uno no puede ser consciente de diferentes estados al mismo tiempo, de la misma forma que uno no es consciente a la vez de tener cinco, diez y ochenta años, aunque pueda tener recuerdos de esas etapas por separado. Pero entonces ¿cómo se determina cuál de esas personas somos? En principio, seremos conscientes de la posibilidad más probable: por ejemplo, es posible atravesar el suelo gracias

a algo llamado el efecto túnel, pero las probabilidades de ello son prácticamente nulas. Según la teoría de Everett, existe una copia de nosotros que ha experimentado el efecto túnel, aunque la inmensa mayoría de nuestros «yos» nunca verán nada parecido. Si aceptamos que somos una de nuestras múltiples copias elegida al azar, la probabilidad de ver el efecto túnel es prácticamente cero: necesitaríamos ser todas ellas a la vez.

Sin embargo, el razonamiento anterior puede llevarnos a la conclusión errónea de que solo somos conscientes de ser una de las múltiples copias o de que nuestras copias no son conscientes de sí mismas. Esto no es así: todas y cada una de nuestras copias son conscientes de existir, aunque solo podamos experimentar una de ellas a la vez.

Uno puede también preguntarse por qué no puede la consciencia existir en una superposición de estados, como un electrón. La respuesta es que el cerebro es un órgano constituido por muchas partes y muy caliente para estándares microscópicos, lo que destruye la coherencia casi de inmediato, eliminando toda posibilidad de interferencia. Esto sugiere que nuestro cerebro se rige, a todos los efectos, por las leyes de la física clásica. Sin embargo, algunos autores como el físico y matemático inglés Roger Penrose (n. 1931) sostienen que nuestras mentes realizan operaciones matemáticas que no pueden explicarse usando la física clásica, así que tiene que darse algún tipo de fenómeno cuántico.

Un físico es solo una manera que tiene un átomo para mirarse a sí mismo.

NIELS BOHR

EL BORRADOR CUÁNTICO DE ELECCIÓN RETARDADA

Si hay un experimento que logra ilustrar la mecánica cuántica de forma aún más espectacular que el de la doble rendija, este es el «borrador cuántico de elección retardada». Se trata de una variante de la doble rendija donde la información sobre por cual de ellas ha pasado la partícula se obtiene después, no antes, de

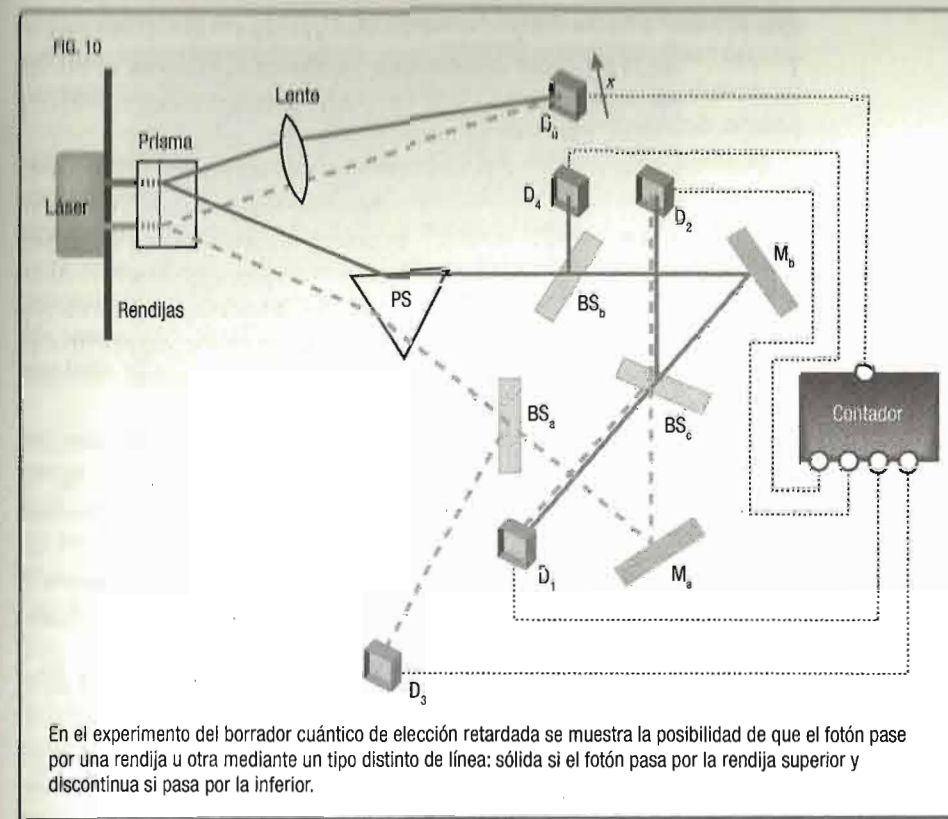
que la partícula quede registrada en la pantalla. A pesar de eso, el patrón de interferencia desaparece igualmente, como si el hecho de averiguar la rendija en cuestión pudiera tener un efecto sobre otra medida realizada en el pasado.

Este experimento, realizado por primera vez en 1999, suele citarse a veces como un ejemplo de «retrocausalidad», es decir, de un proceso donde la causa precede al efecto en el tiempo. Algunos autores argumentan además que el experimento hace imposible una interpretación realista de la mecánica cuántica y que la única salida es aceptar que el observador determina la realidad. Como se verá, ambas afirmaciones son falsas, aunque en primera instancia pueda parecer lo contrario.

A pesar de su aparente complejidad, el dispositivo experimental (figura 10) se basa en una idea relativamente sencilla. En lugar de electrones, se usan fotones, que son las partículas que constituyen la luz. Estos son, en primer lugar, emitidos por un láser y luego pasan a través de una doble rendija. Supongamos que el láser manda los fotones de uno en uno: entonces, cada fotón pasará por ambas rendijas a no ser que se ponga un detector en una de ellas. Es decir: en principio el fotón pasa por ambas si no es medido.

Después de la doble rendija se introduce un prisma especial que divide cada uno de los fotones en otros dos fotones idénticos con la mitad de la energía que el original. Estos dos fotones están entrelazados, ya que provienen del fotón anterior. Así pues, si por ejemplo descubrimos que el fotón sólido que se dirige hacia abajo ha pasado por la rendija superior, sabremos que su pareja de arriba lo ha hecho también. De esta forma, podemos obtener la información sobre el camino que ha seguido el fotón sin tener que poner un detector en la rendija.

Como puede verse, los dos fotones superiores pasan por la lente e impactan contra una pantalla en D_0 , donde son detectados. Dado que los dos fotones son de hecho el mismo, interferirán entre sí y crearán en la pantalla un patrón de interferencia, a no ser que se detecte por qué rendija han pasado, en cuyo caso se destruirá la coherencia y se obtendrá un patrón como en el caso de los balines.



En la zona inferior, los dos fotones pasan por otro prisma PS, donde son desviados hacia los divisores de haz BS_a y BS_b . Estos tienen una probabilidad del 50% de reflejar el fotón y otro 50% de transmitirlo. Fijémonos en el fotón sólido: si es reflejado por BS_b , será dirigido al detector D_4 . Dado que a este solo puede llegar el fotón sólido, que es el que ha pasado por la rendija superior, sabremos inmediatamente por dónde ha pasado también el fotón que ha llegado a D_0 y, por lo tanto, la coherencia quedará destruida y no tendremos un patrón de interferencia. Si, por el contrario, el fotón sólido es transmitido a través de BS_b , se reflejará en el espejo M_b e irá a parar al detector D_1 o D_2 , al que también irá a parar el fotón discontinuo. Dado que cualquiera de los

dos fotones podría haber recalado en D_1 o D_3 , no podemos saber de qué fotón se trata y, por lo tanto, no sabremos por qué camino ha pasado el fotón de D_0 ; en ese caso, tendríamos que obtener un patrón de interferencia.

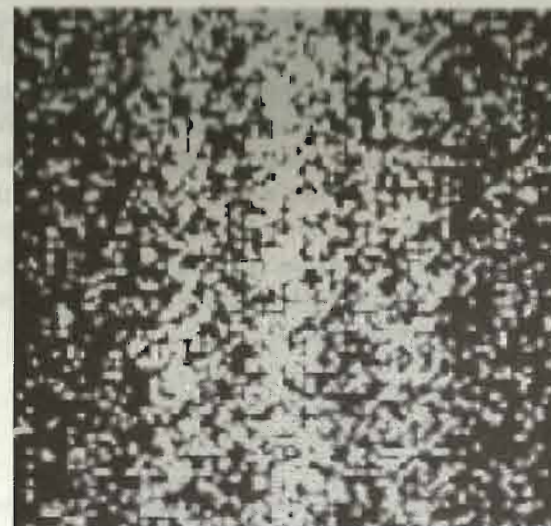
El resultado de todo este dispositivo experimental es que tenemos un 50% de posibilidades de averiguar por dónde ha pasado un fotón, 8 nanosegundos después de que haya impactado contra una pantalla. Si no sabemos por dónde ha pasado, tendríamos que ver un patrón de interferencia; si lo averiguamos, el patrón de interferencia tendría que desaparecer, a pesar de que la información sobre por qué rendija ha pasado se obtiene después del impacto contra la pantalla.

En efecto, si uno selecciona solo los fotones para los que no se conoce el camino que han seguido, obtiene el patrón de interferencia de la figura 11, mientras que, si selecciona aquellos para los que se conoce el camino, se obtiene el patrón que aparece en la figura 12. Es decir, una vez averiguamos por dónde ha pasado el fotón, la interferencia queda destruida, a pesar de que el patrón se formó antes de saberlo.

Este experimento ha llevado a algunos filósofos a afirmar que esto demuestra que nuestra consciencia determina la realidad y no a la inversa. Dado que la elección de un observador consciente afecta a lo que observa incluso después de que lo haya observado ¿qué otra explicación podría haber? Sin embargo, basta con aplicar las reglas de la mecánica con cuidado para ver que eso no es así.

Centrémonos de nuevo en los dos fotones sólidos, recordando que están entrelazados. Cuando el fotón superior es absorbido por la pantalla, ambos quedan entrelazados. Sin embargo, el entrelazamiento no solo se da entre la pantalla y el fotón superior: ¡se da también con el inferior! Dado que ambos fotones estaban entrelazados entre sí, el estado del aparato de medida tiene que quedar entrelazado con el fotón sólido que aún no se ha detectado y cuya función de onda incluye aún dos posibilidades: en una, la información sobre el camino se pierde, mientras que en la otra se descubre. En la interpretación de múltiples universos, diríamos que nuestro sistema vive en una superposición de todas las posibilidades y que estas no quedan completamente

FIG. 11



Patrón de interferencias obtenido a partir de los fotones detectados en D_0 y D_1 .

FIG. 12



Patrón obtenido a partir de los fotones detectados en D_0 y D_4 .

determinadas hasta que realizamos la segunda medida. Si nos ceñimos a lo planteado en Copenhague, encontrar un fotón en una región incompatible con la interferencia elimina la posibilidad de detectar su camino en un futuro: las dos medidas poseen una correlación, pero no una relación causal de presente a pasado. Además, no tenemos que explicar nada, ya que negamos que nuestras ecuaciones se refieran a una realidad subyacente: las ecuaciones no mienten y predicen la desaparición del patrón de interferencia, aunque la información sobre el camino se obtenga con retraso. Lo que está pasando realmente es una cuestión que deberían afrontar los filósofos.

CIENCIA, NO PSEUDOCIENCIA

Se ha visto que la mecánica cuántica es un conjunto de herramientas matemáticas que hacen predicciones experimentales. Esas herramientas tienen problemas filosóficos que pueden ser subsanados mediante una interpretación, es decir: un aparato conceptual adicional que no modifica las predicciones de la teoría.

Pero existe cierta confusión entre interpretación y ciencia y eso hace que, en algunos libros pseudocientíficos, la mecánica cuántica sea citada como la fuente de afirmaciones dudosas sobre la naturaleza, como que «el observador determina lo que ve» o que «la realidad es subjetiva». Se suele afirmar entonces que todas esas conclusiones se han demostrado científicamente, dando una pátina de respetabilidad a lo que el autor quiere demostrar.

Sin embargo y como se ha visto, un mismo fenómeno puede ser visto de formas muy distintas, sin que las predicciones experimentales cambien lo más mínimo. Si bien es cierto que, sea la que sea la interpretación escogida, uno se ve obligado a abandonar ciertas creencias sobre el mundo, también es cierto que esas creencias no son siempre las mismas. Al fin y al cabo, las interpretaciones de la mecánica cuántica y, con ellas, su visión de la consciencia, parecen ser una cuestión más de gusto que de ciencia. Hasta que se encuentre una forma de distinguirlas experimentalmente, permanecerán en el ámbito de la filosofía.

Tiempo y consciencia en la relatividad de Einstein

La mecánica cuántica y la teoría de la relatividad especial de Einstein revolucionaron la física del siglo xx. Ambas parecían ir en contra del sentido común: mientras que la primera socavaba la idea de un universo determinista, la relatividad puso en duda el conocimiento de la época sobre el espacio y el tiempo.

En 1887 Albert Abraham Michelson (1852-1931) y Edward Morley (1838-1923) descubrieron que la velocidad de la luz es una constante, la mida quien la mida, gracias a un célebre experimento que tumbó la teoría del éter. Fue un resultado que nadie esperaba, a pesar de que era una predicción incontestable del electromagnetismo postulado por James Clerk Maxwell (1831-1879), uno de los grandes éxitos de la física del siglo XIX.

En 1864 este físico británico propuso, a partir de observaciones y experimentos anteriores, la teoría electromagnética moderna, gracias a la cual dos fenómenos presuntamente distintos, la electricidad y el magnetismo, fueron unificados en un gran cuerpo teórico que daba cuenta de todo lo que se sabía sobre cargas y corrientes eléctricas. Se apoyó especialmente en el trabajo de su compatriota, el físico y químico Michael Faraday (1791-1867), quien descubrió que un campo magnético oscilante producía un campo eléctrico.

Al analizar las predicciones de su teoría, Maxwell descubrió que una carga en movimiento tenía que emitir ondas: estas consistían en una oscilación en el espacio de los campos eléctrico y magnético. Cuando Maxwell calculó la velocidad de esas ondas

vio con sorpresa que su velocidad era exactamente la de la luz. De ahí concluyó que la luz tenía que ser, pues, una onda electromagnética.

Sin embargo, los contemporáneos de Maxwell pronto se encontraron con un rompecabezas: la teoría electromagnética predecía que la velocidad de la luz era de 299 792 458 m/s, pero no especificaba respecto a qué. Toda velocidad se mide relativa a algo: por ejemplo, una mesa está quieta respecto al suelo, pero se mueve a gran velocidad respecto al Sol. Para una persona que está sentada en un coche, el volante está quieto, pero para otra que se ha detenido en la acera, el volante se mueve a la velocidad del coche. Una velocidad no tiene sentido si no se especifica en qué sistema de referencia la estamos midiendo: es decir, cómo se mueve la persona o dispositivo que la mide.

Así, el valor de la velocidad de la luz obtenido por Maxwell abría una serie de interrogantes. ¿Qué observador la mediría? Parece lógico que, si nos acercamos a una fuente de luz, observaremos una velocidad de la luz mayor que si nos alejamos de ella, de la misma forma que dos coches que chocan a 100 km/h cada uno tienen una velocidad relativa de 200 km/h. Es un contrasentido pensar que el otro coche se acercará a nosotros a 100 km/h sin importar cómo nos movamos hacia él.

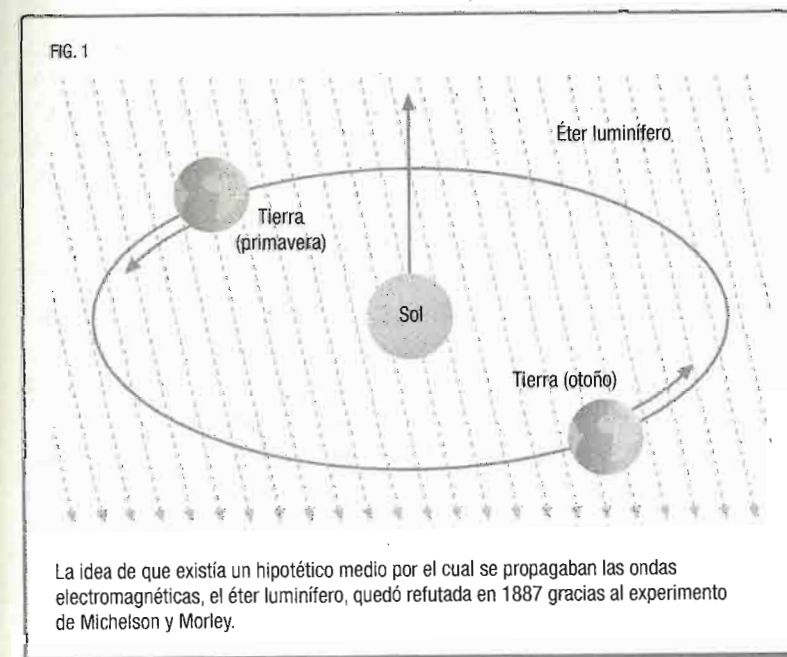
¿POR DÓNDE SE PROPAGAN LAS ONDAS?

En el clima intelectual de la época resultaba claro que no podía ser que la velocidad de la luz que daban las ecuaciones de Maxwell funcionase para cualquier observador, se moviera como se moviera. Sin embargo, las ecuaciones no daban indicación alguna sobre qué observador debía tomarse como referencia. Por suerte, lo que se conocía entonces sobre las ondas dio una pista a la comunidad científica.

Toda onda necesita un medio para propagarse: el sonido lo hace por el aire y las olas por el agua. Una conjetura razonable era pensar que la luz también necesitaba un medio para ser transmitida. Dado que la luz podía propagarse del Sol a la Tierra sin

que hubiese nada entre ellos, los físicos pronto decidieron que tenía que haber algo indetectable en el espacio, un medio por el cual la luz pudiera propagarse. Lo llamaron *éter luminífero* y solucionaba dos problemas: por un lado, proporcionaba un medio para la luz, que de otra forma se comportaría como ninguna otra onda; por otro, daba una respuesta a quién medía la velocidad de la luz: tenía que tratarse de un observador en reposo respecto al éter. Según esta teoría del éter, la Tierra se mueve alrededor del Sol y, por consiguiente, a través de dicha sustancia (figura 1). Por ello, si medimos la velocidad de la luz en diferentes direcciones, tendría que haber una diferencia apreciable que dependerá de la velocidad que la Tierra tenga respecto al éter.

Pero Michelson y Morley determinaron, en su famoso experimento de 1887, que no existía tal diferencia. La luz tiene siempre la misma velocidad, la mida quien la mida y se mueva como se mueva. Se trata de un resultado que atenta contra el sentido común: lo que nos dice es que, tanto si nos acercamos a una fuente



de luz como si nos alejamos de ella, siempre mediremos la misma velocidad de la luz respecto a nosotros. Usando la metáfora de los coches de nuevo, es como si hubiera un coche que siempre avanzase hacia nosotros a 100 km/h, aunque nos alejásemos de él a 70 km/h o nos acercásemos a 100 km/h.

El rompecabezas sobre la velocidad de la luz permaneció abierto hasta que Einstein, en 1905, propuso su teoría de la relatividad especial. En ella, en lugar de intentar explicar por qué la luz se mueve siempre a la misma velocidad, simplemente asumió la realidad de este hecho y pasó a deducir sus consecuencias en las leyes de la física. Pero estas resultaron ser aún más difíciles de digerir que la afirmación que las originaba.

La teoría de Einstein se puede resumir en dos postulados:

1. Cualquier observador que se mueva en línea recta y a velocidad constante observará las mismas leyes de la física.
2. La velocidad de la luz medida por cualquier observador es siempre la misma.

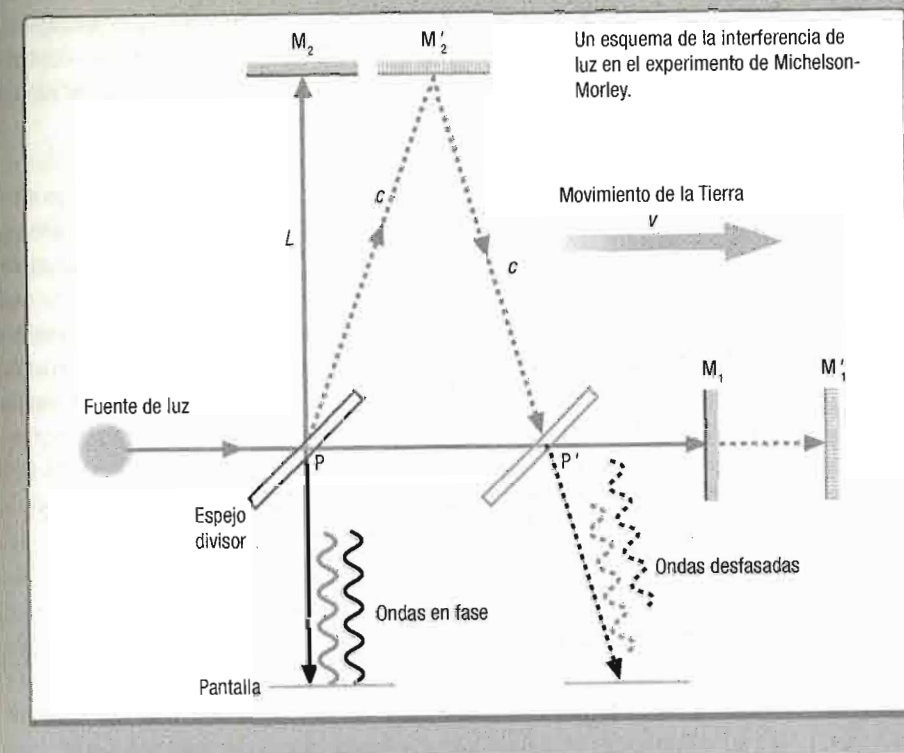
El número 1 es una afirmación que se da sin prueba, a partir del cual se deduce el número 2 y el resto de la teoría. La validez de los postulados se decide en base a su acuerdo con el experimento. De entrada ambos parecen razonables: el primero es necesario si queremos tener un universo consistente, donde las leyes de la física no cambien de un lugar al otro. El segundo surge de la repetida observación de que la velocidad de la luz nunca cambia. Sin embargo, al combinarlos se llega a una serie de conclusiones que parecen absurdas pero que han sido comprobadas una y otra vez en el laboratorio. Hasta la fecha, todas las predicciones de la relatividad han sido confirmadas.

EL TIEMPO ES RELATIVO

Quizá la predicción de la relatividad que más difícil resulta de aceptar es la llamada *dilatación temporal*. Según esta, dos ob-

EL EXPERIMENTO DE MICHELSON-MORLEY

El célebre experimento se basa en la interferencia entre dos haces de luz. Cuando ambos se juntan en un mismo punto, interfieren de forma destructiva o constructiva, dependiendo de la diferencia en el camino recorrido por cada onda. Michelson y Morley se basaron en este hecho para encontrar hipotéticas diferencias en la velocidad de la luz, usando un instrumento llamado *interferómetro*. El experimento, ilustrado en la figura, se inicia en una fuente de luz que crea el haz que se quiere estudiar. Luego se divide ese haz en dos mediante un divisor, enviando la mitad a un espejo superior, M_2 , y la otra mitad a otro espejo colocado enfrente y a la misma distancia, M_1 . Ambos son reflejados de vuelta y se unen en una pantalla situada debajo de dicho divisor. Imaginemos que nuestro dispositivo experimental está quieto respecto al éter: en ese caso, la luz tardará lo mismo en reflejarse en ambos espejos, por lo que el camino recorrido por ambos haces será el mismo, dando lugar a una interferencia constructiva. Por el contrario, si nuestro dispositivo experimental tiene una cierta velocidad con respecto al éter, al hallarse la Tierra en movimiento, no observaríamos la interferencia constructiva. Así pues, estudiando el patrón de interferencia de la luz puede saberse si la velocidad de la luz cambia con el movimiento terrestre o si, por el contrario, permanece constante sin importar cómo se mueva la Tierra.



servadores que se muevan a velocidades distintas medirán distintos intervalos de tiempo para un mismo suceso. Para ver por qué, basta fijarse en un sistema de dos espejos entre los cuales se hace rebotar un fotón, que es como se llama la partícula de luz (figura 2).

La velocidad de la luz es una de las más importantes constantes fundamentales de la naturaleza.

ALBERT ABRAHAM MICHELSON

Imaginemos que tenemos dos espejos paralelos, entre los cuales hacemos rebotar un fotón. A su vez, tenemos un sistema idéntico, a la derecha, pero que se mueve respecto al primero a una cierta velocidad. De lo que hay que darse cuenta en primer

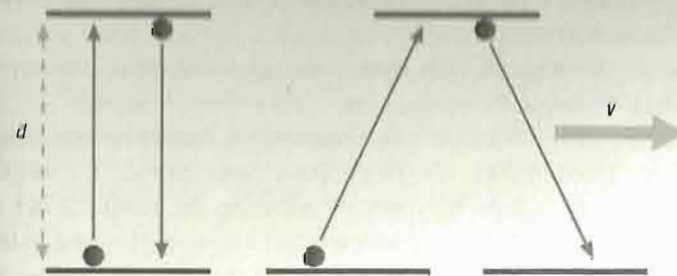
lugar es que esta situación es exactamente la misma si el espejo de la derecha está quieto y el que se mueve es el de la izquierda (figura 3), que es justo lo que verá un observador que se mueva con el espejo de la derecha en la figura 2. Ambas figuras representan la misma situación, vista por dos observadores distintos. En la primera, el observador se mueve con el espejo de la izquierda y en la segunda, con el de la derecha.

Dado que el fotón es una partícula de luz, se mueve a la velocidad de la luz y, de acuerdo con el segundo postulado, ambos observadores deberían medir la misma velocidad. De hecho, la velocidad de la luz es lo único en lo que ambos van a poder ponerse de acuerdo. Veamos por qué. Fijémonos primero en el espejo de la izquierda de la figura 2. Supongamos que los dos espejos están separados por una distancia d . En este caso, si llamamos c a la velocidad de la luz, el tiempo que le tomará al fotón ir de un lado al otro será:

$$t = \frac{d}{c}.$$

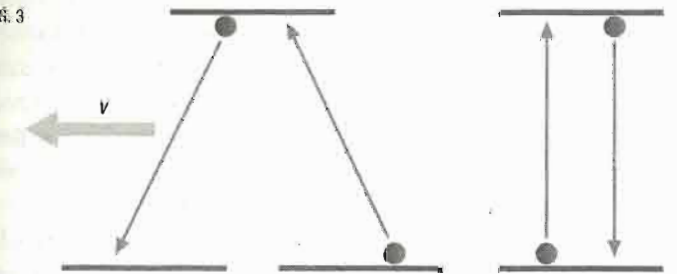
Veamos ahora lo que el observador de la izquierda medirá para el fotón de la derecha: en este caso, la distancia a recorrer es mayor porque, aunque la separación entre espejos sigue siendo d , el fotón se mueve en diagonal. Dado que la distancia es

FIG. 2



Un fotón rebotando entre dos espejos. El conjunto de la derecha se mueve con respecto al de la izquierda.

FIG. 3



El mismo grupo de espejos de la figura 2, visto por el observador de la derecha.

mayor y que la velocidad tiene que ser la misma, el tiempo que transcurre entre que el fotón sale del espejo inferior y golpea el superior tiene que ser mayor que en el primer caso. Es decir: el observador de la izquierda dirá que el fotón de la derecha tarda más que el suyo en llegar al otro espejo.

Sin embargo, el observador de la derecha dirá lo contrario. Según él, es su fotón el que avanza en línea recta, mientras que el de la izquierda lo hace en diagonal. Así pues, concluirá que el tiempo que su fotón se toma para ir de un lado al otro es menor que el del conjunto de la izquierda.

La discusión anterior revela una conclusión sorprendente: ambos observadores medirán tiempos distintos para un mismo suceso. El observador de la izquierda verá que el tiempo trans-

curre más lentamente para el de la derecha, ya que el mismo suceso toma más tiempo. De la misma forma, el observador de la derecha dirá lo contrario.

¿Quién tiene razón? En este caso, tenemos que atenernos al primer postulado: cualquier observador que se mueva en línea recta y a velocidad constante observará las mismas leyes de la física. Así pues, ambos tienen razón: a pesar de que no están de acuerdo en el tiempo medido, sí están de acuerdo en la velocidad del fotón. Tenemos que aceptar entonces que, a pesar de que la velocidad de la luz no es relativa, porque todo el mundo mide la misma, el tiempo sí lo es: dos observadores medirán diferentes intervalos de tiempo para un mismo suceso, diga lo que diga el sentido común.

Esta predicción aparentemente absurda ha sido comprobada una y otra vez en el laboratorio. Por ejemplo, tomando dos relojes atómicos y dejando uno en el suelo y poniendo el otro en un avión y viendo cómo, al aterrizar el avión, la sincronización se ha perdido. La relatividad del tiempo es, además, la única forma de explicar cómo es que partículas con una vida media de microsegundos lleguen a nuestro planeta intactas, a pesar de viajar durante lo que parecen años para un observador en la Tierra.

UN MARCO NECESARIO: EL ESPACIO-TIEMPO

Siguiendo el espíritu del siglo xx, nos volvemos a encontrar con una teoría que explora la relación entre el observador y lo observado. En mecánica cuántica, el observador parece determinar lo que observa por el mero hecho de realizar una medida; en la relatividad, es el movimiento del observador lo que condiciona lo que mide. Esto, de nuevo, entra en conflicto con el paradigma anterior, que asumía que el universo tiene que aparecer igual a quien sea que lo mida. El hecho de que una película dure dos horas no tendría que depender de cómo se mueva nadie. Sin embargo, esto es lo que pasa: para un observador que se mueve a una velocidad cercana a la de la luz, la película parecerá mucho más lenta.

DEDUCCIÓN DE LA DILATACIÓN TEMPORAL

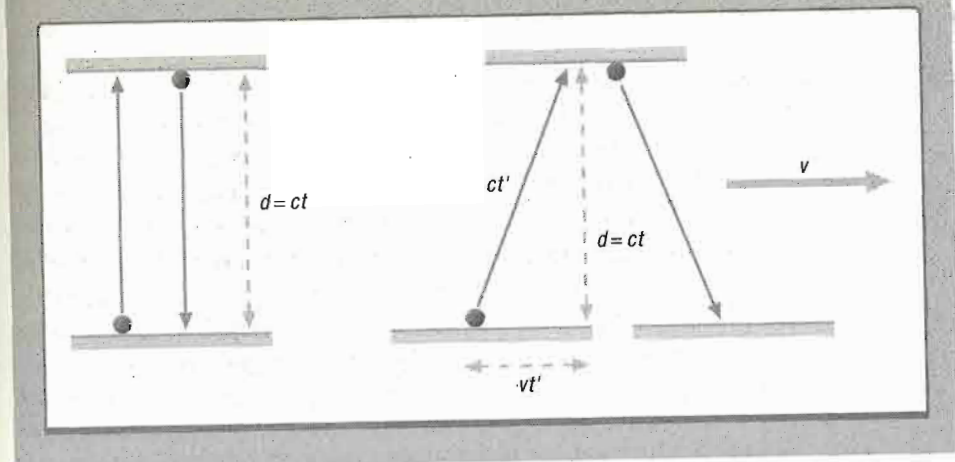
El esquema de los dos espejos y el fotón puede usarse para deducir la fórmula de la dilatación temporal. Para ello tenemos que calcular la distancia que atraviesa el fotón en el caso del espejo en movimiento, como muestra la figura. El observador de la izquierda mide un tiempo t para el tiempo que le toma a su fotón ir de un lado a otro del espejo, mientras que mide un tiempo t' para el fotón de la derecha, ya que este se está moviendo en diagonal y tarda más en llegar. Para obtener la fórmula relativista de la dilatación temporal, basta considerar la longitud de la diagonal del triángulo de la derecha. Como puede verse, su longitud es la de la hipotenusa de un triángulo con lados ct y vt' . Usando el teorema de Pitágoras, obtenemos que la diagonal tiene la siguiente longitud: $(ct')^2 = (vt')^2 + (ct)^2$. Reorganizando términos, tenemos: $(ct')^2 - (vt')^2 = (ct)^2$. Si ahora agrupamos los factores de t' , llegamos a la siguiente expresión: $(c^2 - v^2)t'^2 = (ct)^2$. Dividiendo por $(c^2 - v^2)$ llegamos a:

$$t'^2 = \frac{(ct)^2}{(c^2 - v^2)}$$

Dividiendo todo por c^2 y tomando la raíz cuadrada, llegamos a la fórmula de la dilatación temporal:

$$t' = \frac{t}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Si nos fijamos en la fórmula, veremos que el denominador no puede ser nunca mayor que uno, ya que ninguna velocidad puede ser mayor que la de la luz. Cuando mayor sea v/c , menor será el denominador y mayor, por lo tanto, la dilatación temporal.



El efecto de la relatividad temporal en nuestra vida cotidiana es prácticamente nulo, porque se necesitan velocidades próximas a las de la luz para que sea apreciable. Sin embargo, si el ser humano logra algún día dejar la Tierra y viajar a otros sistemas solares, el efecto tendrá que ser tenido en cuenta. Si una persona viaja a gran velocidad hasta Próxima Centauri, la estrella más cercana, y luego regresa, se encontrará con que el tiempo que ha pasado en la Tierra es mucho mayor que el que ha experimentado: es posible que todas las personas a las que conociese hayan fallecido ya.

Sin embargo, en este caso no hablamos del tiempo psicológico, sino del tiempo real. No es que sintamos que el tiempo pase más rápido o más lento: de hecho, una persona viajando a una velocidad próxima a la de la luz experimentará un paso del tiempo igual al de otra persona que está quieta, porque el movimiento es relativo. Simplemente, los dos observadores no podrán ponerse de acuerdo sobre cuánto dura una determinada acción. Así, parece que el universo cambie de aspecto según quién y cómo lo observa: como si los observadores, por el hecho de estarse moviendo, distorsionaran la realidad. Como se verá, ese no es el caso: la realidad es la misma para todos, pero para verlo necesitaremos de un nuevo constructo que engloba dos entes anteriormente separados: el *espacio-tiempo*.

CONTRACCIÓN DE LONGITUDES Y ESPACIO-TIEMPO

Siguiendo un razonamiento parecido al de la dilatación temporal, es posible deducir que la longitud de los cuerpos sufre un cambio parecido: de la misma forma que el tiempo parece pasar más lento para objetos que se mueven a una velocidad próxima a la de la luz, su longitud se contrae.

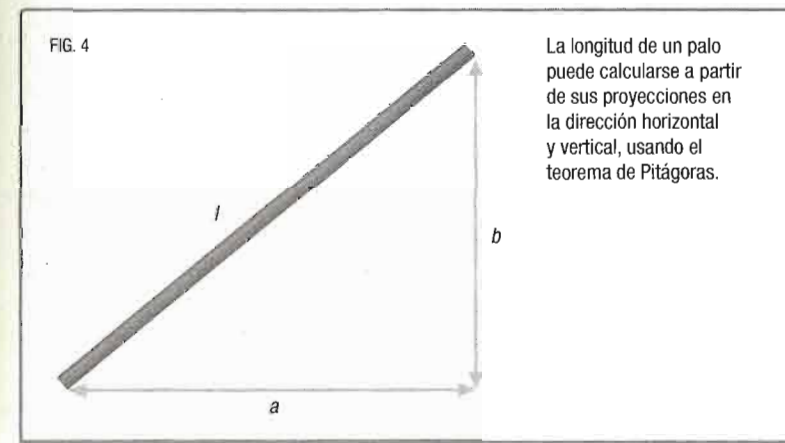
Este descubrimiento, que en un principio parece un problema añadido, nos da una pista sobre la estructura misma del espacio. Mientras que los intervalos temporales se dilatan, los espaciales se contraen. ¿Podría ser que esos dos fenómenos estuviesen relacionados? ¿Que se compensasen de alguna forma?

Esta reducción de una coordenada e incremento de otra se da en otro caso muy común en matemáticas: la rotación. Supongamos que tenemos un palo de dos metros de longitud y lo ponemos sobre una mesa. Por supuesto, podemos medir directamente la longitud del palo. Pero hay otra forma de hacerlo: dividir la mesa en dos direcciones, que por conveniencia llamaremos «vertical» y «horizontal» y medir la distancia horizontal y vertical entre sus dos extremos (figura 4). Llamemos a la distancia horizontal a y a la vertical b . Si queremos encontrar la longitud del palo l , tenemos que usar el teorema de Pitágoras:

$$l^2 = a^2 + b^2.$$

Supongamos por un momento que somos un ser unidimensional, que solo puede ver una dimensión del palo, pongamos que la horizontal. Para nosotros, la longitud del palo será solo a , no l . Si un humano decidiera girar el palo treinta grados, tendríamos la situación de la figura 5.

Pensemos ahora en lo que vería el ser unidimensional: para este, ¡el palo parecería haberse contraído! Sin embargo, la contracción no sería real, ya que el palo seguiría teniendo una longitud de dos metros. Simplemente, el ser unidimensional no se-



ría consciente de ello, porque no tendría acceso a una de las dimensiones.

¿Qué pasa si el ser unidimensional tiene también una noción de tiempo? A cada instante, solo puede ver lo que está inmediatamente delante de él. Sin embargo, cada cierto tiempo se mueve, al margen de su voluntad, una cierta distancia hacia arriba, como se muestra en la figura 6.

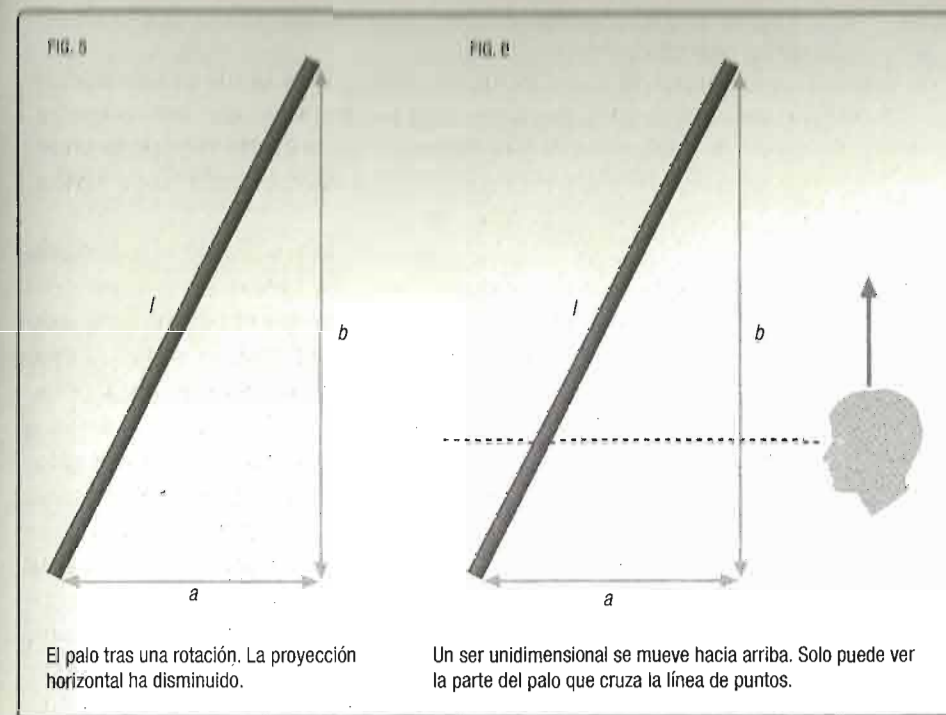
Si los hechos no encajan
con la teoría, cambie los hechos.

ALBERT EINSTEIN

En este caso, al llegar al palo vería que tarda un cierto tiempo en cruzarlo. Ese tiempo estará relacionado con la longitud real del palo, aunque el ser unidimensional no será consciente de ello. Si ahora volvemos a rotar ese palo, podremos crear diferentes efectos. Por ejemplo, si lo ponemos de forma horizontal, el ser lo contemplará en toda su longitud pero lo cruzará en un instante; si lo ponemos de forma casi vertical, lo verá muy contraído pero tardará mucho más en cruzarlo. Si el ser unidimensional tuviera una mente científica, pronto se daría cuenta de que parece haber una relación entre la longitud del palo que observa y el tiempo que tarda en cruzarlo. Tras suficiente deliberación, encontraría esa longitud que nunca cambia en función de la longitud observada y el tiempo para cruzarlo.

Los humanos nos parecemos a ese ser unidimensional: por un lado, tenemos tres dimensiones espaciales por las que nos podemos mover libremente; por el otro, hay otra dimensión, el tiempo, donde solo podemos observar un instante a la vez. De la misma forma que el ser unidimensional, podemos deducir que los objetos que medimos deben tener una longitud independiente de nuestro movimiento, pero que solo tenemos acceso a la parte de longitud confinada a nuestro espacio.

De ahí nace la idea del *espacio-tiempo*: una estructura de cuatro dimensiones que engloba tanto el tiempo como el espacio. Nosotros solo tenemos acceso a una rodaja de espacio-tiempo a cada instante: a esa rodaja la llamamos «espacio». Sin embargo, la relatividad nos muestra que no podemos separar espacio y tiempo, sino que ambos forman parte de un mismo ente.



El descubrimiento de que, a pesar de lo que creemos, no tenemos acceso a todo el «espacio» sino que estamos confinados a una «loncha» infinitamente fina, pone de manifiesto lo poco equipado que está nuestro aparato mental para entender el universo. Por un lado, nuestra intuición nos da la respuesta equivocada cuando las situaciones que consideramos se apartan de lo familiar, como al considerar velocidades próximas a las de la luz. Por el otro, ni siquiera podemos ver el universo en su totalidad y, por si fuera poco, nuestros modelos mentales del espacio y el tiempo están en conflicto con la realidad. Como dijo Einstein, «lo más incomprensible del universo es que es comprensible». Nuestros cerebros están adaptados a sobrevivir en el entorno de la Tierra, no a descubrir los misterios del cosmos. Lo que resulta increíble es que, a pesar de tenerlo todo en contra, incluyendo una intuición que sistemáticamente nos lleva por el

DEDUCCIÓN DE LA CONTRACCIÓN DE LONGITUDES

Consideremos que, tal y como muestra la figura, tenemos un autobús que se mueve entre dos puntos A y B separados por una cierta distancia d , medida por un observador en reposo respecto a A y B. Tenemos, de nuevo, dos observadores: uno de ellos estará en el autobús, mientras que el otro estará en el suelo. Nos planteamos ahora la situación siguiente: cuando el autobús pasa el punto A, tanto el observador de su interior como el que está en tierra ponen en marcha sus cronómetros, que detienen cuando el vehículo pasa el punto B.

La percepción desde la quietud y el movimiento

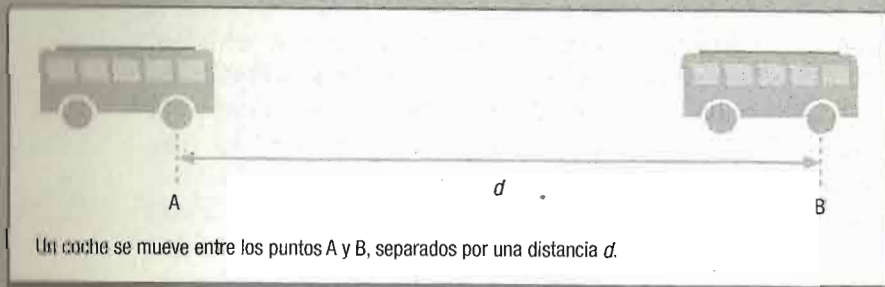
Fijémonos ahora en lo que observa el observador del autobús: según él, el cronómetro del observador que está quieto va más lento que el suyo, porque experimenta dilatación temporal. Por lo tanto, podemos concluir que existe la siguiente relación entre el tiempo t_1 medido por el observador en tierra y el t_2 medido por el que va en el autobús:

$$t_1 = \frac{t_2}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Según el observador en tierra, la distancia entre A y B se puede obtener multiplicando la velocidad del autobús por el tiempo que tarda en cruzar, es decir: $d_1 = vt_1$, donde v es la velocidad del vehículo. Según el observador que está en el interior del mismo, la distancia será su velocidad por el tiempo que él mide: $d_2 = vt_2$. Usando la relación entre t_1 y t_2 que hemos encontrado antes, obtenemos:

$$d_2 = vt_2 \sqrt{1 - \frac{v^2}{c^2}} = d_1 \sqrt{1 - \frac{v^2}{c^2}}$$

Como vemos, cuanto mayor sea v , menor será d_2 : la longitud parece contraerse para el observador en movimiento.



camino equivocado, hayamos podido entender tanto. Y es que, aunque nuestra limitada consciencia suele caer en ciertas trampas, la razón es capaz de sortearlas, lo que nos permite plantear teorías para explicar fenómenos que nuestros ancestros ni siquiera habían imaginado.

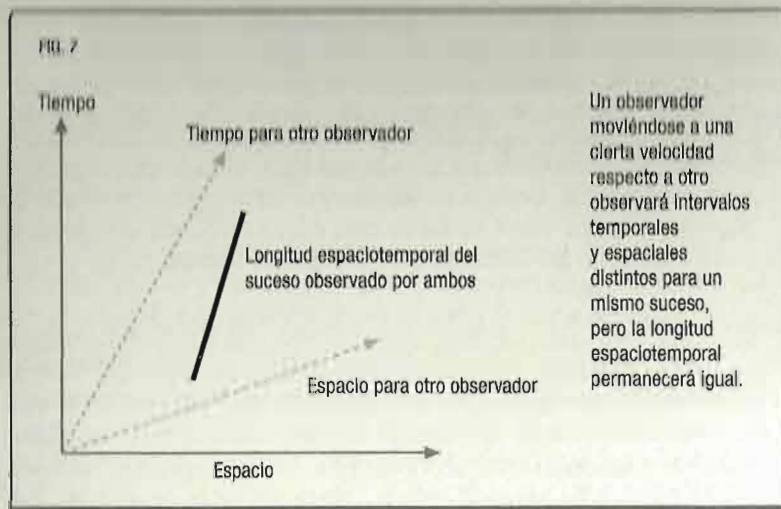
Un poco más allá: juntando el espacio con el tiempo

Podemos desarrollar más la idea de espacio-tiempo si nos damos cuenta que, de la misma forma en que rotar el palo producía una contracción de longitudes y una dilatación temporal para nuestro ser unidimensional, un cambio de velocidades tiene un efecto parecido para nosotros: la longitud se contrae pero el tiempo pasa más lento. Sin embargo, los dos efectos se cancelan para dar lugar a una nueva «longitud» a la que llamamos «intervalo espaciotemporal» y que nos da la separación entre dos puntos en el espacio-tiempo. De la misma forma que podemos utilizar el teorema de Pitágoras para calcular la longitud del palo, podemos usar una fórmula parecida para calcular intervalos espaciotemporales, donde solo consideramos la coordenada horizontal del espacio por simplicidad:

$$s^2 = x^2 - (ct)^2$$

Como puede verse, no se trata exactamente de lo mismo: la coordenada temporal contribuye negativamente a la longitud, lo que hace que nuestra analogía de la rotación no sea completamente adecuada. Sin embargo, las «rotaciones» en el espacio-tiempo tienen una estructura matemática muy parecida a las rotaciones en el sentido usual del término.

Podría parecer que el espacio-tiempo es un invento obvio: al fin y al cabo, todo lo que hemos hecho es tomar el espacio y el tiempo y pegarlos. Esto no es así: lo que el espacio-tiempo nos enseña es que algo que puede parecer «espacio» para un obser-



vador puede a su vez parecer «tiempo» para otro (figura 7). Es decir: el espacio y el tiempo forman parte de una misma estructura y son indisolubles, hasta el punto de que el hecho de que un intervalo sea «espacio» o «tiempo» depende del observador.

Uno puede imaginar el espacio-tiempo como si fuera una serie de diapositivas dispuestas en una torre, y el tiempo sería la altura a la que se encuentra una determinada fotografía. Para asemejar más la imagen a la realidad, debemos considerar que hay un número infinito de diapositivas separadas por una distancia infinitamente pequeña, de forma que toda la torre parece una especie de bloque de queso. Ahora bien, de la misma forma en que es posible cortar ese bloque en lonchas horizontales y paralelas al suelo, también lo podemos hacer inclinando el cuchillo en un cierto ángulo, lo que creará una distorsión en las imágenes. Este «inclinarse el cuchillo» es similar a lo que sucede con nuestra visión del espacio cuando cambiamos de velocidad: el espacio se ve distorsionado, porque estamos viendo lonchas de espacio-tiempo con un cierto ángulo respecto a si estuviéramos quietos.

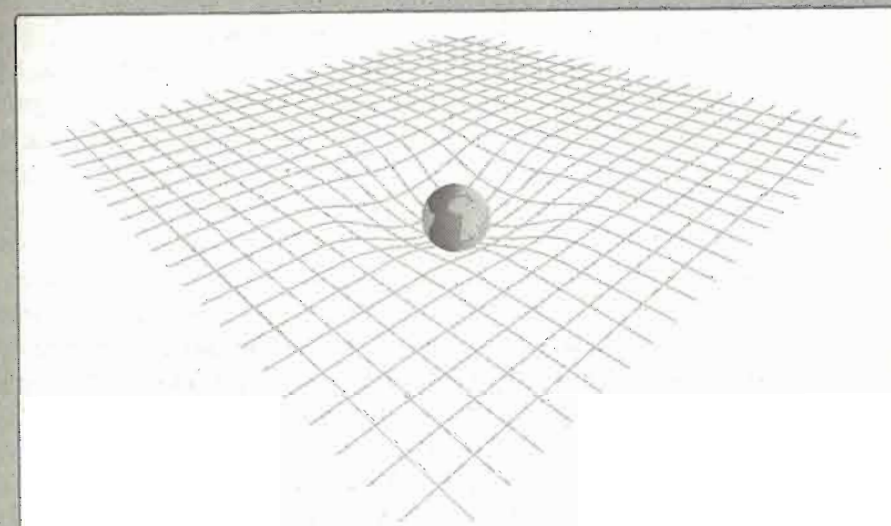
De nuevo, la metáfora de la torre de queso puede llevarse solo hasta un cierto punto: la estructura del espacio-tiempo es más

LA DISTANCIA GENERALIZADA

La introducción de una coordenada negativa al calcular distancias espaciotemporales sugiere otras modificaciones al concepto de distancia. Por ejemplo ¿podríamos tener dos coordenadas negativas, que hicieran las veces de tiempo? ¿Podríamos hacer que una de las direcciones contase más que las otras al medir una distancia? La respuesta es afirmativa en ambos casos. De hecho, es incluso posible crear una nueva definición de distancia que cambie a cada posición del espacio, como, de hecho, hace la relatividad general.

Las curvas del espacio-tiempo

Las distancias generalizadas se usan sobre todo para describir matemáticamente espacios curvos, como la superficie de una esfera. En el caso de la relatividad general, un cuerpo masivo, como ilustra la figura, distorsiona el espacio-tiempo a su alrededor, creando una dilatación del espacio-tiempo que se puede medir. Por ejemplo, si enviamos pulsos de luz a intervalos fijos desde la Tierra a un astronauta en órbita, este verá pasar más tiempo entre dos pulsos que la persona que los está produciendo. Un caso extremo de curvatura espaciotemporal es el de un agujero negro, un cuerpo tan masivo que ni siquiera la luz puede escapar de él. Al acercarnos al centro de un agujero negro el espacio está cada vez más curvado, hasta el extremo de que las distancias se vuelven infinitas. Al punto de máxima curvatura se le llama *singularidad*.



Espacio-tiempo curvo alrededor de la Tierra. La curvatura hace que las longitudes cerca de la Tierra sean mayores que lejos de ella.

sutil y no se adapta a nuestras intuiciones sobre la geometría que vemos cada día, dado que el tiempo actúa como coordenada negativa. De la misma forma, es imposible imaginar una estructura de cuatro dimensiones, así que lo único a lo que podemos aspirar es a crear formas de visualizarla. Lo que sí tienen en común la torre de queso y el espacio-tiempo es el hecho de que la dirección que se considera «tiempo» y la que se considera «espacio» dependen del observador y, por lo tanto, ambas son indisociables.

NUESTRA CONSCIENCIA ACERCA DEL ESPACIO-TIEMPO

La idea del espacio-tiempo tiene consecuencias inevitables en nuestra forma de entender el mundo. Si no tiene sentido decir que el espacio «pasa» de derecha a izquierda, ¿tiene sentido decir que el tiempo «pasa» de pasado a futuro? Ya hemos visto que la dirección considerada como «tiempo» depende del observador: si eso es así, el tiempo tiene que ser exactamente como el espacio y, por lo tanto, no puede estar «pasando» sino que tiene que estar siempre allí. Por supuesto, la palabra «siempre» no tiene sentido en este contexto, porque requiere imaginar un tiempo fuera del espacio-tiempo: más bien podemos ver el espacio-tiempo como una estructura inmutable, congelada, donde todas sus regiones existen simultáneamente, aunque, de nuevo, la palabra «simultáneamente» no debe ser tomada más que como una forma de expresarse.

La idea de que el tiempo no pasa no es solo una consecuencia de la relatividad especial: de hecho, hay poderosos argumentos filosóficos para ella. En concreto, si el tiempo fluyese uno podría determinar a qué ritmo. Pero ¿cómo se mediría eso? Para ello, necesitaríamos un tiempo fuera del tiempo con el que comparar. Otra posibilidad sería decir que el tiempo fluye a un ritmo de un segundo por segundo, pero eso, o es una perogrullada o bien un sinsentido.

Las repercusiones del espacio-tiempo sobre lo que creemos sobre nosotros mismos son difíciles de ignorar. Si el tiempo no «pasa» sino que está siempre allí, ¿por qué sentimos que pasa?

Si el universo no es más que una estructura congelada ¿tiene sentido hablar de acciones, deseos y proyectos? Para responder a esta pregunta, nos ayudará pensar en la película *Desafío total* (*Total Recall*, Paul Verhoeven, 1990).

En ella, Arnold Schwarzenegger es un aburrido trabajador que decide implantarse los recuerdos de unas vacaciones en Marte. Pero algo va mal durante la operación, lo que provoca que el protagonista acabe descubriendo que, de hecho, es un espía en el planeta rojo y que todos sus recuerdos como trabajador en la Tierra son implantados.

La historia de *Desafío total* plantea una cuestión interesante: ¿cómo podemos saber que no estamos en la misma situación que su protagonista? ¿Cómo podemos saber que nuestro pasado es real y no una serie de recuerdos implantados?

La respuesta obvia es «porque yo recuerdo haber ido a la escuela y haber trabajado toda la vida», pero ¿y si esos recuerdos no fueran reales? ¿Notaríamos alguna diferencia?

Podemos llevar este experimento mental más allá: ¿cómo podemos estar seguros de que existíamos hace cinco minutos? ¿Y hace cinco segundos? ¿Y hace una milésima? La respuesta es que no podemos: todo lo que sabemos es el estado actual de la información almacenada en nuestro cerebro, pero no tenemos acceso alguno al pasado. El pasado es una construcción de nuestras mentes, motivada por el hecho de que la información que se encuentra en ellas viene determinada, en principio, por ciertos acontecimientos que se producen en esa dirección temporal y no en la contraria. Todo lo que necesitamos para creer que tenemos un pasado y que el tiempo fluye es un instante, a condición de que la información necesaria exista en nuestro cerebro.

Si ahora volvemos al espacio-tiempo, podemos intentar aplicar lo que hemos aprendido a la pregunta inicial: si el tiempo no pasa, ¿por qué sentimos que transcurre? La respuesta es que solo necesitamos un instante para tener esa sensación: nuestros «yo» de dos, cinco, veinte y ochenta años coexisten y todos y cada

Pero si yo no soy yo... ¿quién demonios soy?

DOUGLAS QUAID (ARNOLD SCHWARZENEGGER
EN *DESAFÍO TOTAL*)

uno creen que el tiempo ha «pasado» hasta llegar al instante actual, después del cual llegará el futuro. Pero esto no es así: en el modelo del espacio-tiempo como estructura fija, todo instante es eterno.

Estamos ante uno de los muchos ejemplos donde la física hace que tengamos que abandonar ciertas intuiciones no solo sobre el universo, sino sobre nosotros mismos. El mismo Einstein consoló así a la viuda de un amigo suyo recién fallecido:

«[...] Se ha ido de este pequeño y extraño mundo un poco antes que yo. Eso no significa nada. La gente como nosotros, que cree en la física, sabe que la distinción entre pasado, presente y futuro es solo una ilusión persistente».

Física y libre albedrío

La visión mecanicista del universo no parece ser compatible con la consciencia de libertad de las personas. La relatividad y la mecánica newtoniana son teorías donde el pasado determina todo instante futuro; en cambio, la mecánica cuántica deja espacio a la aleatoriedad. ¿Somos simples autómatas con una ilusión de libertad? ¿O es la física compatible con el libre albedrío?

La idea de que el libre albedrío es imposible nace al mismo tiempo que la visión determinista del universo. Si el mundo y las personas están hechos de partículas que obedecen una serie de leyes rígidas, su devenir está determinado. Según este paradigma, las personas no son más que máquinas orgánicas: una serie de mecanismos físicos y químicos que dan lugar a ciertos comportamientos. El libre albedrío no es más que un constructo, una ilusión colectiva derivada de la enorme complejidad de los mecanismos que nos gobiernan.

Ante esta tesitura, existen diferentes opciones. La primera es aceptar que el libre albedrío es una ficción y aprender a vivir con ello. Por supuesto, esto tiene consecuencias directas en la convivencia entre personas: si no tenemos libre albedrío ¿existe la responsabilidad moral? ¿Es un psicópata culpable, si solo actúa según lo que le mandan las leyes de la física? El hecho de que las ramificaciones de rechazar el libre albedrío sean incómodas no significa, sin embargo, que la idea sea equivocada. Una de las virtudes de la ciencia es que la aceptación o rechazo de una hipótesis no depende de lo que quisiéramos que fuera cierto.

La segunda opción es negar la mayor: quizá las partículas se comporten siguiendo leyes físicas, pero las personas no lo hacen. A partir de aquí, hay diferentes versiones. La religiosa diría que el alma es incorpórea y que no está sujeta a leyes materiales. Es esta alma la que tiene la potestad de tomar decisiones y afectar el comportamiento del cuerpo. Otra posibilidad es que el mero hecho de tener un gran número de partículas interactuando de forma compleja cree nuevos fenómenos que no son predecibles a partir del simple movimiento de las partículas. Uno de esos fenómenos sería el libre albedrío.

La opción intermedia es no negar que las personas se rijan por las leyes de la física, pero buscar algún resquicio en esas leyes para el libre albedrío. Alternativamente, puede buscarse una definición de libre albedrío que, sin ser incompatible con las leyes físicas, se refiera a algo lo suficientemente parecido a lo que suele entenderse por este concepto.

¿UN LIBRE ALBEDRÍO PREDECIBLE?

Los conflictos entre física y libre albedrío vienen de lejos: surgen en la Grecia clásica, donde los átomos de Demócrito eran incompatibles con la noción de la libertad humana. Sin embargo, la de Demócrito no era una teoría científica y, como tal, no había una cantidad abrumadora de pruebas a su favor. Como muchas de las teorías protocientíficas de los griegos, uno podía creérselas, pero también podía no hacerlo.

La situación se volvió mucho más difícil de sostener con la llegada de la física newtoniana. Esta, como teoría científica, venía avalada por un gran número de datos: podía predecir desde la trayectoria de una bala de cañón hasta la posición de los planetas. Así pues, resultaba muy complicado rechazarla. La teoría de Newton es la primera gran teoría determinista: según esta, los cuerpos se rigen por una serie de leyes matemáticas que determinan por completo su comportamiento. No hay lugar alguno para la ambigüedad y, por lo tanto, para la elección. Quizá la frase que resume mejor el sentir de la época es la cita de Goethe:

«Cuando Werther se rebela contra la naturaleza, lo hace siguiendo las leyes de la naturaleza».

Durante el reinado de la física clásica imperaba la analogía de la persona como autómatas: un ser mecánico hecho de engranajes que se limita a cumplir las funciones para las que ha sido diseñado. Una metáfora más actual es la de un ordenador: una persona recibe estímulos externos, los procesa de acuerdo con ciertos algoritmos que dependen de su genética e información almacenada, a la que llamamos «recuerdos», y luego produce una señal de salida que se corresponde con una determinada acción. Lo que llamamos libre albedrío sería simplemente el proceso que consiste en tomar la información de que se dispone, simular posibles escenarios y elegir el más beneficioso para el individuo según parámetros establecidos previamente. Un psicólogo evolutivo argumentaría que esos parámetros tienen que estar basados en la máxima propagación de los genes: es decir, un individuo tendrá comportamientos que le lleven a maximizar su descendencia.

Contra esta visión mecanicista se suelen esgrimir diferentes argumentos. Por ejemplo, mucha gente hoy en día decide no tener hijos, lo que parece ir en contra de la idea de que las personas somos máquinas programadas para tener descendencia. Contra esto, un psicólogo evolutivo haría notar que la evolución no maximiza la descendencia *per se*, sino comportamientos que tienen el potencial de maximizar esa descendencia en una época determinada. El invento de los anticonceptivos hizo que ciertos comportamientos que maximizaban la descendencia en el pasado hayan perdido relevancia hoy, pero la evolución no ha tenido tiempo aún de corregir el curso. Otra objeción común es que es imposible predecir el comportamiento de un individuo en un instante determinado. Un día entramos en un restaurante y pedimos una ensalada y otro pedimos un bistec, sin que haya un motivo aparente para una decisión u otra. En este caso, un proponente de la teoría mecanicista del libre albedrío diría que las circunstancias nunca son exactamente las mismas y que, dadas exactamente las mismas circunstancias, incluyendo nuestro estado mental, elegiremos siempre lo mismo. El pro-

blema de esta afirmación y de la anterior es que es imposible de comprobar en el laboratorio, ya que no se puede preparar una situación donde tanto las circunstancias como el estado mental del sujeto sean los mismos. Este es uno de los grandes problemas de las ciencias sociales, que se ven obligadas a utilizar leyes probabilísticas dada la dificultad de controlar todas las variables.

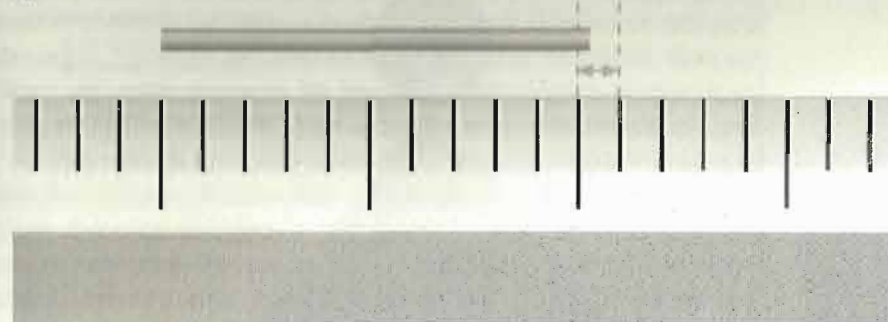
Hoy en día los resultados de diferentes investigaciones en neurociencia parecen apuntar a la posibilidad de que la teoría que niega el libre albedrío apunte en la dirección correcta. Investigaciones recientes demuestran que es posible predecir las decisiones de un sujeto con hasta cuatro segundos de antelación, usando técnicas de resonancia magnética. Parece ser que la toma consciente de decisiones es una ilusión generada por nuestro cerebro, que realiza la mayor parte del proceso de forma inconsciente.

LA INJERENCIA DEL CAOS

Exista o no el libre albedrío, no hay duda alguna de que el comportamiento de las personas es difícil de predecir en el mejor de los casos. ¿Cómo explican eso las leyes de Newton? De hecho, no solo las personas tienen comportamientos erráticos: existe gran cantidad de sistemas físicos cuya evolución nos resulta imposible de predecir. Esto parece entrar en conflicto con la idea de un universo determinado: como suele decirse, si uno conociese la posición y velocidad de todas las partículas en el universo en un cierto instante, debería ser posible predecir su posición y velocidad para todo instante pasado y futuro.

El problema con esta visión es que asume que uno puede medir la posición y velocidad de una partícula con precisión infinita. Sin embargo, esto no es así: toda medida tiene incertidumbre, por muy precisos que seamos. Eso significa que, si conociésemos la posición y velocidad de todas las partículas en el universo, en realidad solo conoceríamos un rango de posibles posiciones y velocidades (figura 1).

FIG. 1

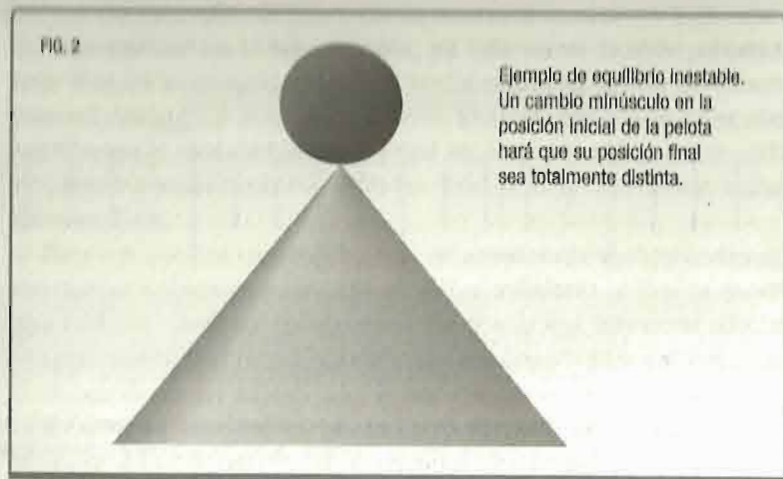


Cualquier medida tiene incertidumbre. En este caso, la incertidumbre en la longitud de la línea viene dada por el hecho de que la regla tiene un número finito de divisiones. La incertidumbre en la posición del extremo derecho es la distancia entre las dos líneas de puntos.

Así pues, si queremos saber el estado del universo en todo instante futuro, no es suficiente con aplicar las leyes de Newton a las partículas con la velocidad y posición que hemos medido: tenemos también que aplicarlas para toda posible combinación de velocidades y posiciones compatibles con nuestra medida.

En un principio podría parecer que eso no es un problema: si las velocidades y posiciones difieren en poco, los estados futuros no pueden ser muy diferentes. Sin embargo, esto no es así: la gran mayoría de sistemas físicos realistas poseen lo que se suele llamar «sensibilidad a las condiciones iniciales»: una pequeñísima variación en el estado inicial del sistema puede tener consecuencias catastróficas en sus estados futuros. A este efecto también se lo conoce como *caos determinista*. Eso se evidencia en el caso de una pelota cuya posición inicial determina el estado final (figura 2). Con que esté tan solo un micrómetro (que equivale a una millonésima de metro) a la derecha, el resultado será completamente distinto a si se hallase un micrómetro a la izquierda.

Existen multitud de sistemas físicos que exhiben sensibilidad a las condiciones iniciales. En general, cuanto más complejo sea el sistema, más impredecible será, aunque existen también sistemas simples con comportamiento caótico. Un ejemplo clásico



sico de sistema caótico es el tiempo meteorológico: su predicción se vuelve más y más difícil a medida que nos alejamos del momento actual y carece de fiabilidad alguna a más de dos semanas vista.

Las personas somos también un sistema caótico: nuestro comportamiento es extremadamente difícil de predecir, al menos usando modelos físicos. Por supuesto, nuestras acciones no pueden ser aleatorias del todo —si lo fueran, ciencias como la economía o la psicología no serían posibles— pero se ajustan mucho menos a principios rígidos que el de un conjunto aislado de partículas simples.

Considerar las personas como un sistema físico caótico arroja luz sobre la cuestión del libre albedrío: quizá su comportamiento sea predecible en principio, pero en la práctica no lo es, aunque se cuente con el instrumental más avanzado y los ordenadores más potentes. A todos los efectos, el comportamiento humano no puede ser reducido al de sus partículas constituyentes. Una reconciliación posible entre física newtoniana y libre albedrío es, entonces, decir que nuestra libertad individual reside en el caos determinista.

Sin embargo, este enfoque no está exento de problemas. Para empezar, parece redefinir la noción de libre albedrío. Según el

caos determinista, nuestro comportamiento sigue estando determinado, pero es imposible de predecir. Eso no es lo que todo el mundo entiende por libre albedrío: por muy difíciles de anticipar que sean, nuestras decisiones siguen sin ser nuestras, sino que vienen determinadas por las leyes de la naturaleza. A esto se podría contraponer que el libre albedrío es precisamente ese proceso determinista pero no predecible que se da cuando tomamos una decisión: la cadena de acciones ocurre dentro de nosotros como máquina orgánica y, por lo tanto, tiene sentido decir que somos nosotros los que hemos decidido una línea de acción u otra. Si bien es cierto que el resultado está predeterminado, eso no nos exime de responsabilidad.

Ninguna de las dos opciones está exenta de problemas y es difícil que satisfaga tanto a quienes defienden que el libre albedrío existe como a quienes creen que es una ilusión. Al final, la ciencia no se pronuncia: la interpretación que cada uno haga del concepto de libertad no es algo que pueda responderse con ecuaciones. La ciencia se limita a hablar de hechos, mientras que la interpretación de estos queda relegada al reino de la filosofía.

LA RELATIVIDAD DEL LIBRE ALBEDRÍO

El problema del libre albedrío se agudiza, si cabe, al considerar la teoría de la relatividad. En esta no solo se mantiene el determinismo de la física newtoniana, sino que el libre albedrío se transforma en algo aún más rígido. En la mecánica de Newton, conocer la posición y velocidad de todas las partículas del universo nos permitiría, en principio, predecir su estado futuro, pero ese estado futuro aún no se ha producido. En la teoría de la relatividad, las leyes de la física se pueden ver como una receta para crear una estructura tetradimensional inmutable. No es que el pasado determine el futuro: es que pasado y futuro existen ya en el espacio-tiempo.

El paso del tiempo es una condición en apariencia indispensable para poder hablar de libre albedrío: hechos en el pasado condi-

clonan una decisión en el presente que tiene consecuencias en el futuro. El libre albedrío habla de elegir entre acciones alternativas pero, si el tiempo no pasa, ¿qué estamos eligiendo? Tanto nuestra decisión como sus consecuencias coexisten en la estructura tetradimensional del espacio-tiempo. Tiene tanto sentido decir que la decisión es anterior a las consecuencias como lo contrario: solo se trata de formas distintas de observar el espacio-tiempo.

Casi todos los físicos del planeta están convencidos de que la relatividad es una teoría correcta pero muchos, a su vez, creen en el libre albedrío. Al fin y al cabo, aunque el universo sea una estructura tetradimensional, los humanos no tenemos acceso a toda ella. En la práctica, nos encontramos en la misma situación que con la mecánica newtoniana la cual, como se ha visto, deja espacio para el libre albedrío siempre que se haga una interpretación lo suficientemente laxa del concepto. Otra forma de encarar el problema es, de nuevo, redefinir el libre albedrío como una función que relaciona estados temporales, sin hacer referencia alguna a si el tiempo pasa o si lo hace de pasado a futuro o a la inversa. Es decir: ciertas características en dos instantes consecutivos de tiempo tienen una cierta relación. Por ejemplo, en un instante una persona está viendo la televisión y quiere cambiar de canal; en otro, hay un cierto canal encendido. La relación matemática que lleva de la primera situación a la segunda es lo que se suele llamar «libre albedrío». Se trata, como en el caso de la física newtoniana, de una solución de compromiso que tiene pocos visos de satisfacer a nadie, pero es el tipo de solución a la que uno se ve abocado cuando tiene que explicar el libre albedrío dentro de los confines de la física.

Sin embargo, aunque el espacio-tiempo como una estructura fija describe muy bien el universo de la relatividad, hay ciertas sutilezas que permiten abrazar una visión más conservadora. Para entenderlas, hay que adentrarse un poco más en los misterios de la *geometría de Minkowski*, que es la que describe la relatividad especial.

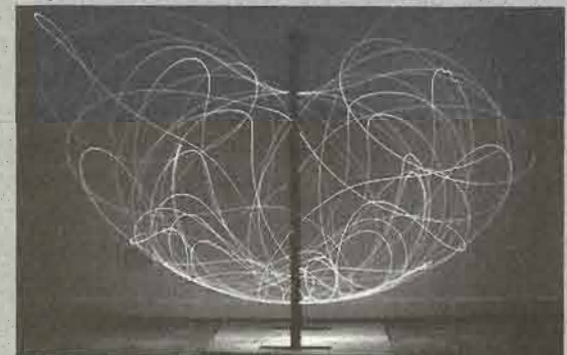
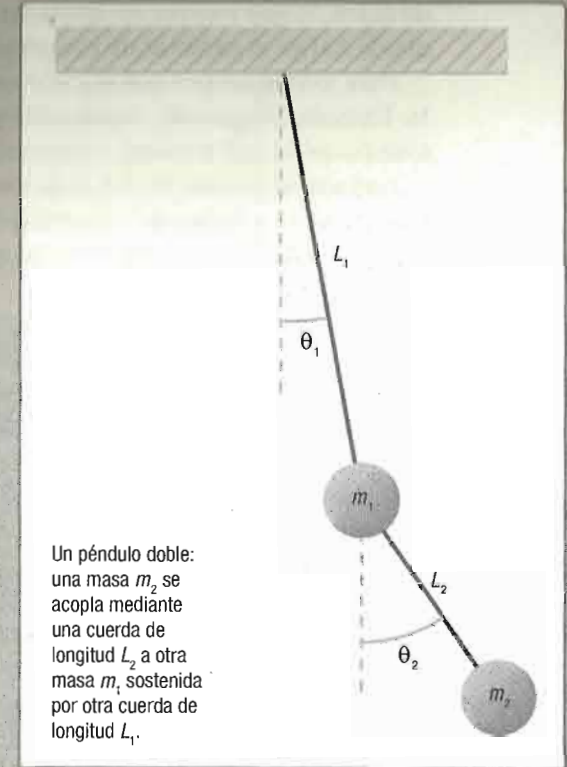
Como se ha visto anteriormente, un cambio de velocidad en relatividad es parecido a una rotación ordinaria en dos dimensiones. Sin embargo, la analogía no es perfecta: al contrario que en una rotación cualquiera, en la relatividad no se puede trans-

EL PÉNDULO CAÓTICO

El caos determinista no se reduce a sistemas complejos, sino que puede darse en condiciones de extrema simplicidad, como en un péndulo caótico: dos péndulos acoplados uno tras otro, tal como se muestra en la figura.

Una trayectoria confusa

El péndulo caótico es un sistema simple y, como tal, relativamente sencillo de modelar. Sin embargo, dadas unas determinadas posiciones iniciales para las dos masas, es imposible predecir dónde se encontrará al cabo de un tiempo. Otra forma de entender esto es fijarse en la trayectoria que ambas masas seguirán en el espacio, por ejemplo tomando una fotografía de larga exposición (en la fotografía aquí reproducida, mediante una luz led se ha trazado la dinámica caótica de un péndulo doble). Aunque se suelte el péndulo desde la misma posición cada vez, se obtendrán fotografías distintas, dado que una diferencia minúscula en las condiciones iniciales acabará teniendo efectos enormes en el futuro. Si algo tan simple exhibe un comportamiento impredecible, podemos imaginar la dificultad de analizar sistemas con billones de partículas, como por ejemplo una persona.



formar por completo un intervalo espacial en uno temporal ni a la inversa. Esto se debe a que no es posible superar la velocidad de la luz, lo que impone un límite en la cantidad de distorsión que puede haber al cambiar de observador.

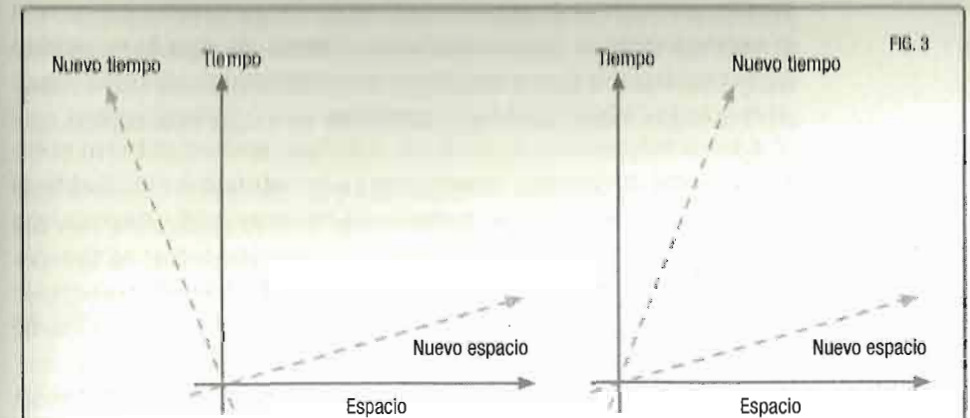
Para averiguar por qué sucede esto usaremos una herramienta llamada *diagrama espaciotemporal*, donde se muestra el tiempo en el eje vertical y una sola dimensión del espacio en el horizontal (figura 3). En la parte izquierda se muestra lo que pasaría si un cambio de velocidad fuese equivalente a una rotación; en la derecha puede verse lo que realmente sucede con los ejes de coordenadas.

Ambos diagramas tienen ciertas características en común: en primer lugar, hay un cambio de ejes de coordenadas, de forma que un mismo punto en el plano será visto de forma distinta. En segundo lugar, la distancia entre dos puntos cualesquiera permanece constante. Sin embargo, hay una diferencia clave: la distancia espaciotemporal puede ser positiva o negativa, mientras que la euclidiana solo puede ser positiva.

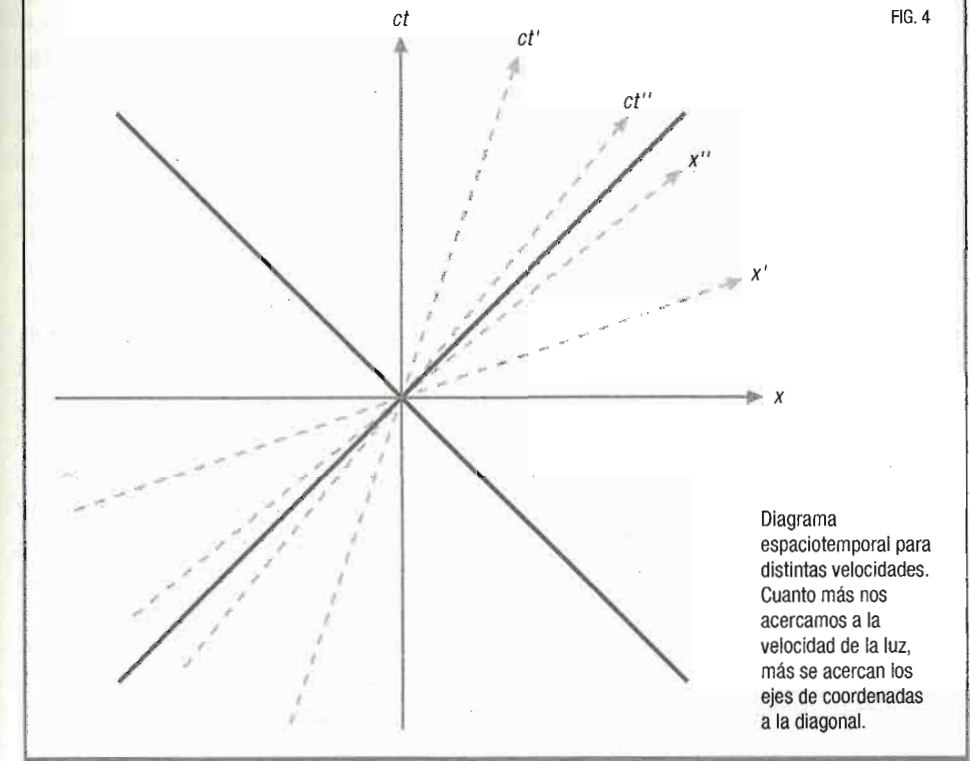
Una distancia espaciotemporal negativa significa que la longitud temporal es mayor que la espacial, ya que la coordenada temporal es la que acarrea el signo negativo. Si la distancia espaciotemporal es positiva, entonces predomina la parte espacial. Ahora bien: dado que la longitud espaciotemporal tiene que ser la misma para todo observador, tiene que suceder que, si la distancia espaciotemporal es negativa para uno, lo sea para todos. Es decir: si el tiempo predomina para un observador, predomina para todos. De la misma forma, si el espacio es dominante para un observador, también lo es para el resto.

Esto significa que podemos clasificar todo intervalo espaciotemporal en dos tipos: *espacial* o *temporal*. Es imposible efectuar un cambio de observador tal que un intervalo espaciotemporal cambie de tipo. Esto podrá apreciarse más fácilmente si vemos cómo cambian los ejes a distintas velocidades (figura 4).

Como puede apreciarse, los cambios de observador están limitados a acercarse más y más a la diagonal derecha, pero lo que nunca puede suceder es que las coordenadas espacial y temporal se inviertan. Es decir: a pesar de que, al cambiar de observador,



A efectos del espacio-tiempo, un cambio de velocidad y una rotación son cosas distintas. A la izquierda se simula que son equivalentes; a la derecha, la diferencia real.



podemos observar modificaciones en la longitud de objetos y en el paso del tiempo, nunca veremos un intervalo espacial volverse temporal. Parece haber una separación fundamental entre espacio y tiempo, dada por la imposibilidad de viajar más rápido que la luz.

La discusión anterior nos muestra que, aunque sí es cierto que el espacio y el tiempo se transforman el uno en el otro, hay un límite a esta transformación. Qué parte de un intervalo es tiempo y qué parte es espacio puede variar, pero la fracción predominante será siempre la misma. Es posible, tensando un poco la cuerda, argumentar que la relatividad no es incompatible con un tiempo que fluye, dado que sí parece haber una diferencia fundamental entre intervalos de tiempo y espacio. Esto deja un poco más de margen a la idea de que el tiempo sí que pasa, que el futuro no está determinado y que, por lo tanto, los humanos gozan de libre albedrío.

El físico y empresario Stephen Wolfram (n. 1959), conocido sobre todo por ser el creador del software *Mathematica*, lleva esta idea un poco más allá. Wolfram cree que el universo es un programa de ordenador y que, como tal, tendría que poder ser generado por un algoritmo que calcula el siguiente estado basándose en el anterior. En este caso, tenemos una noción muy clara de simultaneidad: dos instantes son simultáneos si son generados durante el mismo paso por el algoritmo. La idea de Wolfram es muy parecida a la de Newton: un universo en tres dimensiones que evoluciona en el tiempo según una serie de normas.

Sin embargo, Wolfram ha logrado incorporar la relatividad especial a su sistema, sin abandonar el requerimiento de que cada instante del universo sea generado paso a paso. Lo ha hecho a base de forzar una serie de simetrías en el proceso de creación de su universo, de forma que el constructo generado se parezca al espacio-tiempo de Einstein. A pesar de que las teorías de Wolfram distan mucho de ser aceptadas por la comunidad internacional, su modelo muestra que es posible construir un universo donde el tiempo fluye de acuerdo con la intuición sin desobedecer los postulados de la relatividad. Por lo tanto, la existencia del espacio-tiempo como una estructura tetradimensional inmutable

no es estrictamente necesaria. De la misma forma que uno puede abrazar la forma de pensar relativista y ver el espacio-tiempo como algo real, también puede adoptar una postura más pragmática y considerarlo como una herramienta matemática que describe bien los hechos, pero que no se ajusta necesariamente a la realidad. Los defensores del libre albedrío tienen un resquicio al que aferrarse, sin necesidad de rechazar la ciencia establecida.

LA LIBERTAD CUÁNTICA

Hasta ahora hemos tratado únicamente teorías deterministas como la mecánica newtoniana y la relatividad especial. Sin embargo, la teoría más importante del siglo xx, la mecánica cuántica, no es determinista o, al menos, algunas de sus interpretaciones no lo son. La introducción de una incertidumbre fundamental puede parecer la brecha perfecta para restaurar la idea tradicional de libre albedrío. De hecho, hay gurús de la autoayuda que afirman que la mecánica cuántica abre la posibilidad de que escojamos nuestro destino: dadas todas las posibilidades que nos presenta la mecánica cuántica, podemos escoger la que más se ajuste a nuestro gusto, sin vulnerar ley alguna de la física. Por supuesto, no hay nada en la teoría que permita una interpretación semejante.

Para analizar la relación entre la mecánica cuántica y el libre albedrío es necesario primero escoger una interpretación. La de los múltiples universos y la de Copenhague, a pesar de hacer las mismas predicciones, dibujan una realidad completamente distinta. Mientras que en la interpretación de Copenhague cada medida tiene varias posibilidades entre las que se escoge una aleatoriamente, en la de los múltiples universos el cosmos simplemente se desdobra, de forma que la aleatoriedad es fruto de nuestra percepción, pero no una característica del mundo. Nos centraremos primero en la idea de Copenhague para luego explorar la teoría de los múltiples universos.

Averiguar si la mecánica cuántica en la versión de Copenhague deja espacio para el libre albedrío requiere analizar a fondo

INTERVALOS TEMPORALES Y ESPACIALES

Podemos definir un intervalo *espacial* como aquel en el que, para al menos un observador, la distancia temporal es cero. Imaginemos, por ejemplo, que hay dos personas en los dos extremos de una habitación, encendiendo una linterna al mismo tiempo. En este caso, la separación entre ambas será solo la longitud de la habitación, ya que no hay intervalo temporal alguno: los dos sucesos son simultáneos. Alguien situado en reposo respecto a la habitación construiría un diagrama como ilustra la parte izquierda de la figura 1. La *distancia espaciotemporal* entre los dos sucesos será, entonces, la longitud de la habitación. Ningún cambio de coordenadas puede modificar este hecho, de la misma forma que, sin importar cómo se mire, la longitud de una mesa permanece idéntica. Lo único que puede suceder es que otros observadores vean más o menos parte espacial y temporal, pero la combinación de ambas será siempre la misma.

¿El espacio sobre el tiempo?

Veamos qué pasa si observamos la situación a través de los ojos de otro observador que se mueve a cierta velocidad respecto al primero. En este caso, tendremos un diagrama espaciotemporal parecido al de la parte derecha de la figura 1. Como puede apreciarse, parte de nuestro intervalo espacial se ha transformado en temporal, pero la distancia espaciotemporal permanece igual, lo que significa que el espacio sigue predominando sobre el tiempo.

¿El tiempo sobre el espacio?

Podemos definir los intervalos temporales de forma parecida a los intervalos espaciales: aquellos donde al menos un observador no ve separación espacial alguna. Por ejemplo, una persona chasquea los dedos y luego lo vuelve a hacer diez segundos más tarde. Nuestro diagrama espaciotemporal tendrá el aspecto que se muestra en la figura 2. De nuevo, dado que la distancia espaciotemporal tiene que ser la misma para todo observador, esto significa que el tiempo predominará sobre el espacio, lo mida quien lo mida. Dos observadores pueden estar en desacuerdo sobre cuánta parte es tiempo y cuánta es espacio, pero nunca sobre la distancia espaciotemporal entre dos sucesos. A cantidades como la distancia espaciotemporal se les llama *invariantes relativistas*: son cantidades sobre las que cualquier observador estará de acuerdo. En resumen: los intervalos temporales permanecen temporales para cualquier observador y lo mismo pasa con los espaciales.

¿O ambos iguales?

Hay un tercer tipo de intervalo relativista llamado *intervalo nulo*. En un intervalo nulo la parte de espacio y de tiempo son exactamente iguales, de forma que se cancelan, ya que la parte temporal es negativa. Es decir, la longitud espaciotemporal de un intervalo nulo es cero. Los intervalos nulos son propios de partículas que se mueven a la velocidad de la luz, como los fotones. Dado que la longitud de un intervalo espaciotemporal tiene que ser la misma en todo sistema de referencia, tiene que suceder que, si una partícula se mueve a la velocidad de la luz respecto a un observador, lo hace también respecto al resto.

FIG. 1

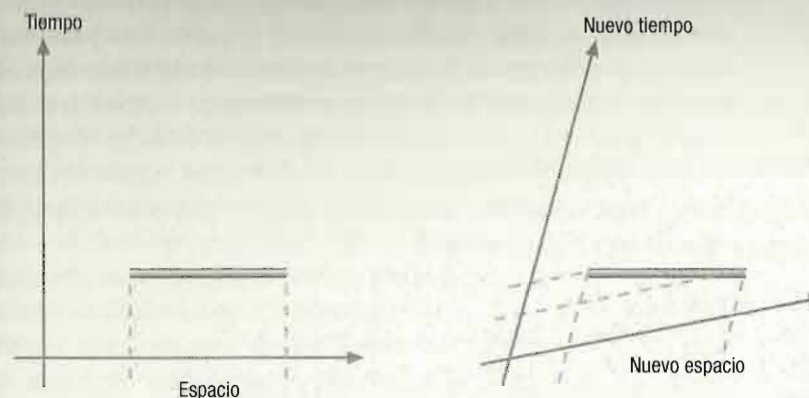


Diagrama que muestra dos puntos separados por una distancia solamente espacial. Las líneas de puntos marcan la distancia espacial entre los dos extremos para un observador en reposo (izquierda) y otro moviéndose a cierta velocidad respecto a estos (derecha).

FIG. 2

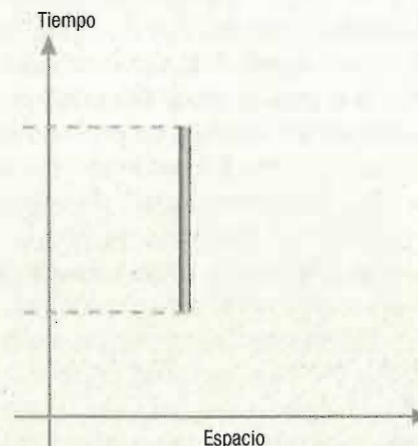


Diagrama que muestra dos puntos separados por un intervalo espacial. Las líneas de puntos marcan la distancia temporal entre los dos extremos.

la propia noción de libertad. En un principio podría parecer que libre albedrío y aleatoriedad están relacionados; dada una cierta situación, se podría escoger hacer tanto una cosa como la contraria y lo que haga es impredecible *a priori*. Sin embargo, la relación no está tan clara. Al fin y al cabo, una persona que tomase decisiones aleatoriamente sería considerada fuera de la norma. La mayoría de la gente actúa como lo hace por alguna razón, salpicada quizá con breves periodos de impredecibilidad. Si la aleatoriedad implica libre albedrío, nos vemos obligados a admitir que un dado o un electrón en un experimento de la doble rendija también lo poseen.

La noción de libre albedrío parece sustentarse en dos pilares. Primero, en la idea de que, a pesar de que acabemos tomando una cierta decisión, podríamos haber tomado otra y nuestra actuación no estaba determinada de antemano. Segundo, en el hecho de que somos nosotros, como entes conscientes, los que decidimos. Pero ¿qué significa eso, exactamente? Este último requerimiento parece estar más basado en una impresión subjetiva que en un hecho contrastable. Cuando tomamos una decisión, somos conscientes de sopesar las diferentes alternativas y de simular los distintos resultados en nuestra cabeza. Todo ese proceso parece ser mental y no físico: al fin y al cabo, la física no hace predicción alguna sobre imágenes mentales. Da la sensación, por lo tanto, de que hay un proceso dentro de lo que interpretamos como nuestra persona que es el que nos lleva a tomar una cierta decisión. Este proceso nos pertenece y, por lo tanto, tendemos a rechazar instintivamente que sea causado por las leyes de la física. Tenemos acceso directo a las causas de la decisión y estas son pensamientos, no colisiones entre partículas.

La interpretación de Copenhague permite el libre albedrío en el primer sentido, pero no en el segundo. Las personas, como todo en el universo, somos sistemas cuánticos y, por lo tanto, nos comportamos según leyes probabilísticas. Dado que nuestros procesos cerebrales se basan, en última instancia, en procesos cuánticos, parece razonable esperar que estén sujetos a una impredecibilidad parecida. Si pensamos en el caos determinista, una diferencia minúscula en un instante dado puede tener con-

secuencias enormes en el futuro; de la misma forma, las diferentes posibilidades que permite la mecánica cuántica para el mundo microscópico pueden traducirse en efectos medibles a gran escala, como nuestro comportamiento.

De la misma forma que un electrón puede pasar por una u otra rendija, una persona puede tomar una u otra decisión. La diferencia entre ambos sistemas es que las personas son objetos macroscópicos y los electrones no. Debido a nuestro tamaño y temperatura, cualquier efecto cuántico tenderá rápidamente a la decoherencia: esto significa que, al contrario del electrón, no podemos hallarnos en una superposición de estados, sino que seremos conscientes de una sola de las muchas posibilidades. Dado que al menos algunas de esas posibilidades se determinan aleatoriamente, nuestra decisión es, al menos en parte, aleatoria también. Por lo tanto, tenemos

razón al entender que, de la misma forma que podríamos haber tomado una determinada decisión, podríamos haber tomado la contraria y que el hecho no estaba determinado de antemano.

La mecánica cuántica de Copenhague sigue fallando en el segundo criterio. Una acción aleatoria no implica voluntad alguna: solo impredecibilidad. Como se ha visto, ambas no son compatibles. No existe ningún mecanismo conocido en el cerebro que sea capaz de escoger entre estados cuánticos: todo lo que sabemos apunta a que, como todo otro sistema, las personas son agregados de partículas que siguen las leyes de la mecánica cuántica. Es decir que, aunque tengamos la sensación de estar escogiendo entre varias posibilidades, en realidad nuestra elección o bien está determinada, en cuyo caso no se trata de libre albedrío, o bien es aleatoria, en cuyo caso tampoco gozamos de libre albedrío.

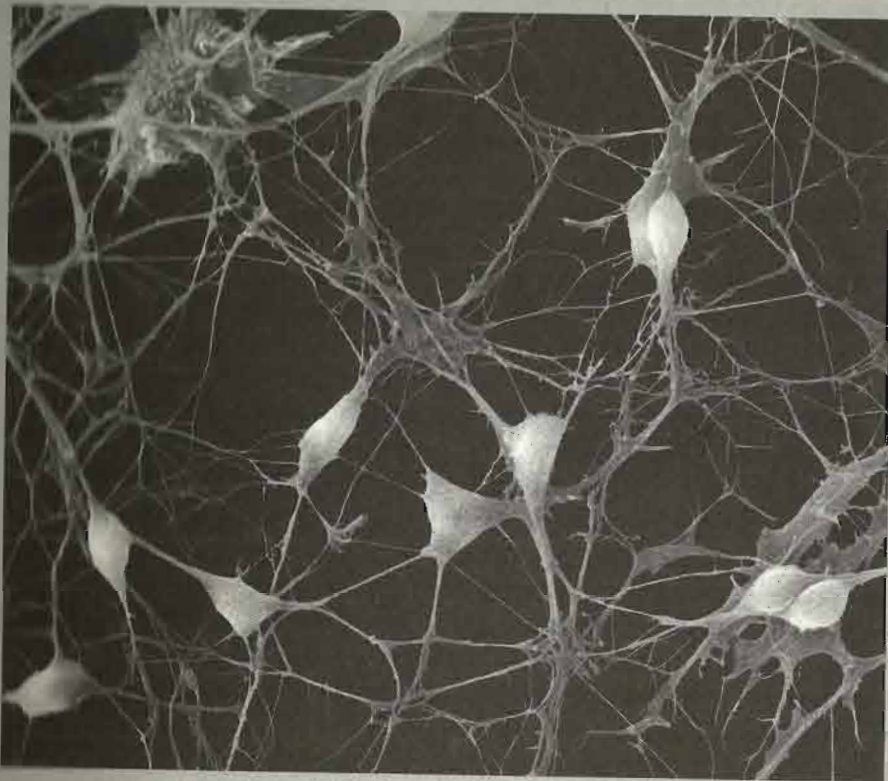
Uno puede intentar salir de esta situación redefiniendo, de nuevo, el significado de libre albedrío. Por ejemplo, podemos

Todas las maravillas de nuestro universo pueden, en efecto, ser capturadas por simples reglas, sin embargo [...] no hay manera de conocer todas las consecuencias de estas normas, excepto observando y viendo cómo se desenvuelven.

STEPHEN WOLFRAM

CONSENSO NEURONAL

Tomar una u otra decisión no es tarea fácil, cómo tampoco lo es analizar por qué optamos por una determinada alternativa. Hasta ahora se creía que, ante una disyuntiva, la confianza en una decisión se dilucidaba entre las dos poblaciones neuronales implicadas, sensorialmente integradas. Pero ahora, un estudio realizado por investigadores del Centro de Cognición y Cerebro de la Universidad Pompeu Fabra de Barcelona y de las universidades Torcuato di Tella de Buenos Aires, Argentina, y Columbia, en Estados Unidos, ha revelado un dato sorprendente: los científicos han comprobado, mediante un modelo computacional, que la decisión final depende de un proceso de «voto» en el que toman parte muchas poblaciones neuronales sensorialmente independientes. La propuesta más votada es la que acabamos por hacer nuestra. Todo un sistema democrático asentado, en su mayor parte, en el interior de nuestro cerebro.



Micrografía electrónica de barrido que muestra numerosas neuronas del córtex cerebral. En la esquina superior izquierda, un astrocito, la más común de las células gliales, sostiene el sistema nervioso.

considerar que lo que nos define como personas son los procesos físicos que ocurren dentro de nosotros. En este caso, afirmar que uno toma libremente una decisión equivale a decir que han sido los procesos físicos internos a la persona los que han llevado a ese resultado. El problema de este enfoque es que parece ir contra la intuición: ¿cómo puede ser que lo que percibimos como deseos e intenciones sea solo el resultado del movimiento de miles de millones de partículas? Este es un problema abierto que, hasta la fecha, no ha sido resuelto. Algunas escuelas filosóficas afirman que no puede ser resuelto y que la consciencia no puede ser explicada en términos físicos; otros afirman que no se trata de un problema en absoluto y niegan que la consciencia pueda ser sujeto de investigación, ya que no es algo que se pueda medir. Los científicos, de momento, se mantienen al margen, aunque existe una creencia generalizada de que los estados mentales tienen que poder ser explicados a través de los estados cerebrales, dado que se ha observado una clara correlación entre ambos. La naturaleza de la explicación sigue, a pesar de todo, eludiendo a la ciencia.

Como antes, nos vemos obligados a escoger entre varias opciones, ninguna de ellas deseable: podemos aceptar que no tenemos libre albedrío porque las personas somos máquinas orgánicas; podemos rechazar que las personas sigan las leyes de la física; finalmente, podemos redefinir lo que queremos decir con «libre albedrío». Cada una de las opciones tiene ventajas e inconvenientes, pero no hay, de momento, ninguna forma de conciliar nuestra intuición con lo que sabemos del universo.

ELEGIR LIBREMENTE EN UN UNIVERSO DESDOBLADO

Puede que la mecánica cuántica de múltiples universos sea, filosóficamente hablando, la más problemática de todas las interpretaciones, en especial en lo que toca a la consciencia. Al contrario que la interpretación de Copenhague, la de los múltiples universos es determinista: dado el estado del universo en un cierto punto, es posible deducir toda su evolución futura, aplicando simplemente la ecuación de Schrödinger a todo el sistema.

Ninguna cuestión es tan sublime como la que se pregunta por qué hay un universo: por qué hay algo en lugar de nada.

DEREK PARFIT

En este caso, al contrario que en la mecánica newtoniana, la evolución del universo no se limita a una sola posibilidad, sino a todos los futuros posibles. Es decir: la teoría predice que todos y cada uno de los futuros compatibles con el presente actual

son reales. Volviendo al libre albedrío, vemos que no solo podríamos haber tomado una decisión distinta a la actual: de hecho, hemos tomado todas y cada una de las decisiones posibles. Lo que pasa es que no somos conscientes de haberlo hecho.

Si uno decide aceptar la existencia de los múltiples universos, se ve abocado a una serie de problemas de índole más fundamental que la simple cuestión de la libertad humana: la definición misma de identidad personal. Las personas somos conscientes de ser el mismo individuo que nuestro yo de hace cinco minutos y distinto que nuestro vecino. El mismo concepto de libre albedrío requiere la existencia de una persona que evolucione en el tiempo y tome decisiones basándose en eventos anteriores. Pero si a cada instante nos desdoblamos en una infinidad de clones que toman decisiones distintas, ¿cuál de ellos somos? ¿Somos todos la misma persona o personas distintas? ¿De cuál de ellas somos conscientes?

Derek Parfit (n. 1942) es un filósofo inglés que ha dedicado buena parte de su vida académica a considerar estas cuestiones. En su libro *Razones y personas* plantea la siguiente situación, que puede ayudar a esclarecer el problema: imaginemos que es el siglo xxx y que alguien ha inventado una máquina de teletransporte. Esta máquina funciona a base de destruir nuestras moléculas en nuestra posición actual y reconstruirlas en otra, por ejemplo en la superficie de Marte. Supongamos que hiciésemos el experimento y entrásemos en la máquina: ¿realmente despertaríamos en Marte? Es decir, la persona reconstruida en el planeta rojo ¿seguiría siendo ella misma?

Es, por supuesto, imposible hacer el experimento. Pero, aunque pudiéramos llevarlo a cabo, seguiríamos sin poder averiguar si esa persona sigue siendo la que era. Supongamos que su

consciencia perece con su cuerpo terrestre: a pesar de todo, hay una nueva persona en Marte que tiene un cuerpo igual al suyo, además de sus recuerdos y personalidad. Si le preguntásemos si el experimento ha funcionado, tendría que decir que sí: al fin y al cabo, recuerda haber entrado en la máquina y despertar en territorio marciano. El clon estaría convencido de ser esa persona. Recordaría haber ido a la escuela, haber trabajado y haberse subido en la máquina en pos de un viaje espacial.

Consideremos ahora un experimento parecido: como antes, una persona sube en la máquina, pero esta vez despierta en la Tierra. Los ingenieros encargados del proyecto le informan, orgullosos, de que han mejorado el teletransportador: ya no hace falta destruir el cuerpo al hacer el teletransporte. Desafortunadamente, el proceso daña el tejido cardíaco, así que el individuo tendrá un ataque al corazón en los próximos diez minutos. Por la pantalla, verá a su clon marciano saludándolo risueño.

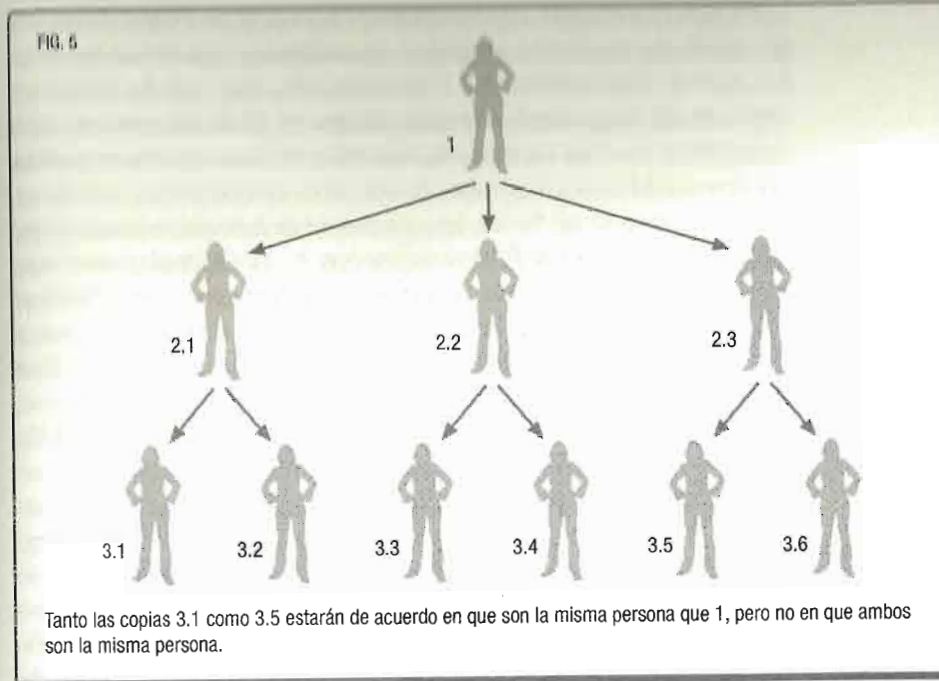
Este experimento mental sirve para sacar a la superficie muchas de nuestras creencias sobre la identidad personal. No tenemos problema alguno en aceptar que hemos sido teletransportados a Marte si nuestro cuerpo es destruido, pero nos cuesta decir lo mismo cuando seguimos vivos en la Tierra, aunque solo sea durante diez minutos. Si todo lo que importa es compartir recuerdos y personalidad, además de una cierta continuidad en nuestra percepción, entonces tendríamos que concluir que tanto nuestro clon como nosotros somos aquella persona que entró en la máquina, aunque ambos afirmaríamos que la otra copia no es realmente «nosotros». De forma alternativa, podemos concluir que el concepto de identidad personal no tiene sentido y que, de hecho, cada una de nuestras instancias personales es una persona distinta, conectada con las otras por una serie de recuerdos y rasgos de personalidad. Según esta visión, el concepto mismo de libre albedrío deja de tener sentido, porque la persona que recaba la información es distinta de la que toma la decisión y de la que recibe sus consecuencias. De hecho, dado que una decisión es un proceso que toma tiempo, resulta muy difícil discernir quién está tomando la decisión y quién está, simplemente, sufriendo sus consecuencias.

En la interpretación de múltiples universos nos vemos con un problema igual al del ejemplo del clon marciano: la persona en cuestión se desdobra a cada microsegundo en miles de clones, cada uno de los cuales tiene continuidad psicológica respecto al anterior. Si preguntásemos a cualquiera de esos clones si son la misma persona que hace cinco minutos, contestarían que sí; si les preguntamos si son la misma persona que ellos mismos en una realidad alternativa, contestarían que no. Pero esto nos deja en una mala posición, porque «ser la misma persona» tendría que ser una propiedad transitiva: si A es la misma persona que B y B es la misma persona que C, entonces A tendría que ser la misma persona que C. Pero ese no es el caso si escuchamos lo que dice nuestra intuición sobre la teoría de múltiples universos (figura 5).

El caso de la mecánica cuántica de múltiples universos se agrava por el hecho de que todas y cada una de nuestras copias tienen el mismo derecho a decir que son la misma persona que nosotros. En el ejemplo de Marte, nuestra intuición nos indicaba claramente que seríamos la misma persona que nuestro «yo» terrestre, no nuestro clon marciano. En el caso de la teoría de múltiples universos, la intuición no sabe qué decir: si nos bifurcamos en veinte personas distintas, todas continuas físicamente con la actual, ¿quién seremos conscientes de ser? Si nos fijamos en cada una de ellas por separado, parece que en cada caso la respuesta es que deberíamos seguir siendo conscientes, lo que nos dice que todas y cada una de nuestras copias son la misma persona desdoblada: deberíamos ser conscientes de ser todas ellas. Sin embargo, no lo somos: solo somos conscientes de ser una persona, no veinte. ¿Pueden conciliarse las dos versiones?

Una posible respuesta es que somos conscientes de todas ellas, pero no a la vez. De la misma forma que, en la relatividad, nuestro yo de cinco años coexiste con el de ochenta y ambos creen ser el único que existe, en la mecánica cuántica de múltiples universos todas y cada una de nuestras versiones son conscientes de ser nosotros, a pesar de que cada una de ellas cree ser la única.

La mecánica cuántica de múltiples universos crea interrogantes difíciles de resolver en la cuestión de la identidad personal, donde nos vemos obligados a replantearnos obviedades apa-



rentes como que solo hay una persona que se corresponde con nosotros. Esto puede llevarse incluso más allá: si todas mis copias tienen el mismo derecho a decir que son la misma persona que yo, ¿qué hay de mi vecino? ¿Qué hace que seamos personas distintas? Este tipo de cuestiones son ineludibles cuando nos vemos expuestos a las dificultades que crea para nuestro sentido común la nueva visión del mundo. Por desgracia, la física no ofrece respuesta alguna, sino que se limita a exponer los hechos. Como hemos dicho, la interpretación es una cuestión de filosofía y, en muchos casos, de preferencia personal.

Dado que es difícil delimitar quién es la misma persona que quién, la noción de libre albedrío sufre de nuevo un problema de falta de sujeto. ¿Quién toma la decisión y quién la sufre? Sin embargo, el problema va mucho más allá: no tomamos una decisión u otra, sino que tomamos toda decisión posible. ¿Cómo se reconcilia eso con la idea de que elegimos libremente?

Según el experto en computación cuántica David Deutsch (n. 1953), la mecánica cuántica de múltiples universos no solo no agrava el problema del libre albedrío, sino que lo resuelve. Deutsch afirma que el culpable de que el libre albedrío no sea compatible con las leyes de la física no es el determinismo, sino el espacio-tiempo de la relatividad: tanto si el universo es determinista como si no lo es, la existencia del espacio-tiempo implica que todo hecho futuro existe ya, lo que significa que uno no puede cambiarlo eliminando por consiguiente la posibilidad del libre albedrío. Incluso si el universo fuera indeterminista y los hechos futuros se escogiesen aleatoriamente, seguiríamos teniendo el mismo problema, porque mientras haya un espacio-tiempo seguirán estando determinados, aunque la forma de determinarlos sea lanzando un dado.

Según Deutsch, la existencia del espacio-tiempo quita el sentido a afirmaciones como «tras pensar las cosas con calma, elegí hacer X; podría haber escogido hacer otra cosa; fue la decisión correcta; se me da bien tomar decisiones de este tipo». Si el futuro está escrito, no existe la posibilidad de hacer otra cosa. De hecho, la noción misma de causa y efecto se desmorona: tiene tanto sentido decir que el presente causa el pasado como que el pasado causa el presente. Dado un instante, todo futuro y pasado están determinados. Como solo hay un futuro posible, tampoco tiene sentido decir que tomamos una buena decisión, porque no tenemos ninguna alternativa para comparar.

En el multiverso, sin embargo, todas esas aseveraciones toman un nuevo significado que es compatible con las leyes de la física, a pesar de que estas son deterministas, al contrario que en la interpretación de Copenhague. Por ejemplo, «tras pensar las cosas con calma, elegí hacer X» puede interpretarse como «tras pensar las cosas con calma, algunas de mis copias, incluyendo la que está hablando, eligieron hacer X». La frase «podría haber escogido hacer otra cosa» deja de ser una imposibilidad lógica y se convierte en «algunas de mis copias han elegido hacer otra cosa». Del mismo modo, «fue la decisión correcta» puede interpretarse como «las copias que han tomado esta decisión son más felices que las que no lo han hecho». «Se me da bien tomar

decisiones de este tipo» puede interpretarse como «hay muchas más copias de mí que toman la decisión que lleva a la máxima felicidad que las que no lo hacen».

Como puede verse, el multiverso permite dar sentido a frases que los filósofos llaman *contrafactuales*: «si esto no hubiera sucedido, habría sucedido lo siguiente». En física clásica y en relatividad no tiene sentido plantearse qué habría pasado si no hubiéramos tomado una determinada decisión: se trata de una imposibilidad lógica. Incluso en la interpretación de Copenhague, donde esas decisiones son aleatorias, mientras aceptemos un espacio-tiempo determinado seguiremos teniendo el mismo problema. En la teoría de múltiples universos, les podemos dar un significado físico: existen copias de nuestro universo donde, de hecho, hemos tomado la decisión contraria. Cuando decimos que, si no hubiésemos decidido ir al cine ese día, no habríamos conocido a nuestra mujer, estamos diciendo que existe un universo donde no elegimos ir al cine y en el que, en efecto, nunca conocimos a la que es nuestra esposa en este universo.

A pesar del entusiasmo de Deutsch, la teoría de múltiples universos es capaz de provocar la reacción opuesta. El cuento de ciencia ficción *All the myriad ways*, de Larry Niven (n. 1938) describe perfectamente por qué. En él, la civilización terrestre ha encontrado una forma de viajar a universos paralelos: al hacerlo, descubren que hay realidades alternativas donde, por ejemplo, los nazis ganaron la Segunda Guerra Mundial. Poco después de que empiecen los viajes comerciales a realidades paralelas, empieza una ola de crímenes y suicidios, que el detective Trimble tiene que investigar.

Pronto se da cuenta de que no se trata de crímenes comunes: en general, las víctimas son gente educada y con dinero que ha viajado hace poco a una realidad alternativa. Trimble tiene la teoría de que algún tipo de virus tiene que haber cruzado la barrera entre mundos y causado la epidemia. Sin embargo, pron-

Sentirse insignificante porque el universo es inmenso tiene la misma lógica que sentirse inadaptado por no ser una vaca.

DAVID DEUTSCH

to se da cuenta de que este no es el caso: parece ser que es el hecho de haber visto las realidades alternativas el que hace que la gente se suicide. Al final del cuento, Trimble se encuentra en su despacho mirando la pistola que tiene encima de la mesa. Y se pregunta qué sucedería si se disparase en la cabeza. Se responde que no pasaría nada: en un sinnúmero de realidades alternativas, seguiría tomándose su café en paz. De hecho, aunque decidiese no disparar, en otro sinnúmero de realidades alternativas ya lo habría hecho. «¿Por qué no?», se pregunta.

El cuento de Larry Niven sirve para ilustrar un aspecto inquietante de la mecánica cuántica de múltiples universos: si ya hemos tomado todas las decisiones posibles, ¿significa algo cualquier decisión que tomemos? Quizá en una realidad paralela no existimos o hemos ganado la lotería. Algunas de nuestras copias tendrán más éxito, otras menos. Pero... si era posible tomar una decisión y la contraria y de hecho hemos tomado esa decisión y la contraria, ¿qué mérito tiene tomar la decisión correcta? ¿Cómo podemos juzgar a una persona por tomar una decisión equivocada, cuando esa misma persona ha tomado la correcta en un universo paralelo?

Como puede verse, la interpretación de múltiples universos resuelve algunos problemas a base de crear otros; la de Copenhague tiene consecuencias parecidas pero opuestas. No deja de resultar chocante que dos teorías que hacen las mismas predicciones tengan interpretaciones filosóficas tan distintas. En lo que coinciden ambas interpretaciones es en la necesidad de revisar el significado de libre albedrío o, en su defecto, rechazar que los humanos se comportan siguiendo las leyes de la física.

EL LIBRE ALBEDRÍO, VERSIÓN TEÓRICA

La relación entre mecánica cuántica y libre albedrío se complicó aún más en 2006, cuando los matemáticos John H. Conway (n. 1937) y Simon B. Kochen (n. 1934) publicaron un teorema matemático al que llamaron *el teorema del libre albedrío*. En este partían de tres premisas basadas en la mecánica cuántica y demos-

traban que, si un físico experimental puede decidir libremente qué propiedades mide en una partícula, entonces el resultado de su medida no está determinado de antemano.

El teorema está basado en tres supuestos o axiomas denominados *Fin*, *Spin* y *Twin*. Estos supuestos se basan en aspectos tanto de la mecánica cuántica como de la relatividad especial de Einstein, aunque son más generales y, por lo tanto, valdrían para una clase más amplia de teorías. Explicados de forma simplificada, los supuestos que usan Conway y Kochen son los siguientes:

1. Axioma *Fin*: la información no se puede transmitir instantáneamente de un sitio a otro. Esto se inspira en el hecho de que la velocidad de la luz es finita, pero funcionaría para otras velocidades.
2. Axioma *Spin*: si medimos el espín de una partícula determinada y lo elevamos al cuadrado, los resultados de nuestra medida serán uno: uno y cero en diferentes órdenes. Esto se basa en experimentos clásicos en mecánica cuántica donde se mide el espín de dos partículas entrelazadas.
3. Axioma *Twin*: es posible entrelazar dos partículas de forma que las medidas de sus espines estén relacionadas.

Como puede verse, los tres axiomas tienen una inspiración clara en la mecánica cuántica, aunque son algo más generales. Las interpretaciones de este teorema son tan variadas como las opiniones que existen sobre la existencia del libre albedrío: defensores de este último aseguran que, tras el teorema, hay una prueba matemática de que las personas poseemos libre albedrío; sus detractores aseguran que lo que Conway y Kochen han demostrado es simplemente que el universo no puede ser determinista si sigue las leyes de la mecánica cuántica, cosa que ya sabíamos.

Para adquirir una mejor idea del significado del teorema y sus repercusiones en la filosofía del libre albedrío, debemos fijarnos primero en lo que Conway y Kochen quieren decir con una elección «libre». Según ellos, una persona toma una decisión libre si

su acción no está determinada por ningún hecho pasado; es decir, no hay información alguna que podamos extraer del pasado que nos pueda permitir predecir esa decisión. En otras palabras, cuando realizamos una medida libre, aparece nueva información en el universo: una información que antes no estaba, porque en caso contrario hubiera sido posible predecir lo que el físico experimental iba a medir.

El teorema de Conway y Kochen implica, por lo tanto, que si la decisión del físico experimental sobre qué medir no depende del pasado, tampoco lo hace la propiedad que se mida de la partícula. Es decir, no hay nada en el pasado que nos permita descubrir qué valor obtendremos de la medida o, en otras palabras, se ha creado nueva información en el universo concerniente al valor de nuestra medida.

Según sus autores, su teorema puede usarse para demostrar que las personas tenemos libre albedrío: dado que las acciones de las partículas no son predecibles ni en principio, tampoco lo son las nuestras. Por lo tanto, nuestro futuro no está determinado y poseemos libre albedrío. Como suele suceder en física, es fácil ponerse de acuerdo sobre el significado matemático y experimental del teorema, pero su interpretación dista mucho de ser unívoca. En este caso, Conway y Kochen asocian el libre albedrío al hecho de tener conductas aleatorias pero, como se ha visto antes, un comportamiento aleatorio dista mucho de una conducta libre. De hecho, prácticamente nadie en la comunidad física discute que el resultado de una medida sea aleatorio y, por lo tanto, imposible de determinar previamente. Así, el teorema del libre albedrío puede verse más como el último clavo en el ataúd determinista que como una demostración de que exista el libre albedrío.

MÁS ALLÁ DE LA FÍSICA SE HALLA LA METAFÍSICA

Si el grueso de la discusión anterior parece más metafísica que física, es porque lo es. La física se ocupa de sucesos simples con definiciones rigurosas: la velocidad de una partícula es algo que se puede medir y que tiene un significado claro que nadie discu-

te. El cometido de la física es encontrar una serie de leyes que den cuenta de todas las relaciones posibles entre esos sucesos simples. Del mismo modo, toda pregunta en física tiene que ser reducible al resultado de un experimento: si no, más que física, se la considera metafísica, que procede de un vocablo griego que significa justo eso: «más allá de la física». Esta forma de trabajar tiene muchas ventajas. Los resultados en física son contrastables y fáciles de comunicar; los principios descubiertos tienen aplicabilidad universal. La limitación a preguntas que pueden responderse en el laboratorio evita que proliferen teorías especulativas, lo que posibilita un avance constante en la comprensión del mundo.

Por desgracia, todo lo que hace a la física una herramienta perfecta para estudiar el movimiento de los cuerpos, la convierte en poco idónea para analizar conceptos más complejos y peor definidos como el libre albedrío o la consciencia. No todo el mundo entiende el concepto de «energía cinética», pero todo el mundo tiene una idea de a qué se refiere cuando dice que toma sus propias decisiones. Sin embargo, resulta difícil en extremo definir con rigor a qué nos referimos. Incluso si nos pusiéramos todos de acuerdo en una definición de libre albedrío, seguiríamos teniendo el problema de que es imposible diseñar un experimento que dé o quite la razón a cierto grupo, porque algunas teorías sobre el libre albedrío son acientíficas: es imposible demostrar que no existe un alma incorpórea ligada al cerebro que no tiene otro efecto físico que dar consciencia y libertad a una persona, porque la teoría del alma no realiza ninguna predicción que pueda ser falsada en el laboratorio.

Eso no significa que el del libre albedrío no sea un problema digno de ser estudiado. Se trata de una cuestión central para organizar la convivencia humana que no puede ser ignorada sin

La gente piensa que las matemáticas son complicadas, pero son la parte simple [...]. Los gatos sí que son complicados. ¿Qué hay en las pequeñas moléculas y otras materias para que un gato se comporte de manera diferente que otro? ¿Qué forma un gato? Y ¿cómo se define un gato? No tengo ni idea.

JOHN HORTON CONWAY

más. La física quizá no pueda darnos respuestas, pero sí puede ayudarnos a encontrar mejores preguntas: ¿son lo mismo libertad y aleatoriedad?, ¿cómo definimos al individuo que toma la decisión?, ¿existe el libre albedrío si no pasa el tiempo?, ¿cómo podemos saber que dos personas son la misma? Todas estas cuestiones surgen de considerar que la física, lejos de ser una ciencia cerrada en sí misma, es un conocimiento cuyas consecuencias van mucho más allá de las partículas elementales a las que describe.

Al contrario que en una empresa científica, la cuestión del libre albedrío está sujeta a las ambigüedades del lenguaje y a los prejuicios de cada uno. No se trata de una pregunta que pueda responderse solo con la ciencia, pero sí se trata de una empresa donde la ciencia puede ayudarnos a buscar una respuesta mejor. No tiene sentido exigir que la filosofía se rija por los parámetros de la física, pero sí puede tomarla como inspiración, construyendo teorías que no entren en conflicto con la experiencia y definiendo sus conceptos con el máximo rigor posible, como ya hace la filosofía analítica. Por desgracia, no hay una forma universalmente aceptada de escoger entre teorías, excepto juzgando los méritos respectivos de sus argumentos. Al final, la creencia o no en el libre albedrío depende de preferencias personales y de un juicio subjetivo sobre lo convincente que sea un determinado argumento.

Cómo simular la mente

Hasta inicios del siglo xx se creía que los fenómenos mentales y los físicos no estaban relacionados. Pero hoy en día la mayoría de científicos cree que la consciencia tendría que poder ser explicada por la ciencia. Si la mente es material, ¿es posible simularla en un ordenador? Y si es así, ¿qué significaría eso para nuestra comprensión de la naturaleza humana?

En su libro *The Mind's Eye*, el físico, matemático y filósofo estadounidense Douglas R. Hofstadter (n. 1945) propone la siguiente cuestión: ¿cómo podríamos detectar que estamos dentro de una simulación perfecta? Hofstadter no se refiere a seres humanos conectados a una máquina, como en *Matrix*, sino a seres humanos simulados dentro de una máquina: un universo donde todo, tanto edificios y partículas como personas, son unos y ceros.

La respuesta es, por supuesto, que no podríamos. Si la simulación es perfecta, cualquier experimento que hagamos dará los mismos resultados que los que realicemos en la realidad. El comportamiento de cada una de nuestras partículas virtuales sería el mismo que el de nuestras partículas reales y no habría diferencia alguna entre el universo simulado y el real. Pero, si una simulación perfecta y la realidad son indistinguibles, ¿tiene sentido tratarlas de forma diferente?

Como el ejemplo de Hofstadter parte de la premisa de que una simulación perfecta es posible, no podemos usarla como argumento a la hora de plantearnos simular una consciencia. Pero plantea una cuestión central: ¿es posible reducir la consciencia a unos y ceros? O, de forma más general: ¿es posible reducir la

consciencia a un conjunto de normas que se aplican siempre? Si la respuesta es afirmativa, entonces la consciencia no es más que información, y el soporte físico que usa, en este caso nuestro cuerpo, es irrelevante. Si no lo es, puede que haya algo en la consciencia que no pueda ser reducido a la mera aplicación de las leyes de la física.

ARGUMENTOS CONTRA EL REDUCCIONISMO

A la teoría de que todo lo que sucede en el universo puede ser reducido a interacciones entre partículas se la suele llamar *reduccionismo*. Aplicado a la consciencia, el reduccionismo afirma que nuestras ideas, recuerdos y sensaciones no son más que configuraciones de partículas, aunque nosotros los percibamos de forma distinta. Dado que las interacciones físicas son simulables, también lo es la mente. Es decir: si aceptamos que las leyes de la física no permiten excepciones, no tenemos más remedio que aceptar también que nuestra mente puede en principio ser simulada. Sin embargo, hay científicos reduccionistas, como Roger Penrose, que, aunque aceptan que la mente se rige por las leyes de la física, no admiten que pueda ser simulada por un ordenador clásico. Es decir: uno que funcione según los mismos principios que el típico portátil de sobremesa.

Si es posible simular la consciencia tiene que ser posible, también, crear máquinas conscientes. La forma más obvia sería «recrear» una persona, de la que ya sabemos que es consciente: si la simulación se comporta de la misma manera, tendremos que concluir que nuestro ordenador también lo es. Sin embargo, tendría que haber una forma más general para poder determinar si una máquina es consciente de sí misma.

En 1950, el matemático británico Alan Turing (1912-1954) propuso una solución. En un artículo, sugirió que una conversación por escrito debería bastar: si un ordenador es capaz de engañar a su interlocutor y hacerle creer que es humano, habrá que concluir que es consciente. Esta idea se conoce como el *test de Turing*. Hasta hoy, ningún ordenador ha logrado pasar un test de Turing

auténtico, aunque algunos programas han sido capaces de convencer de su condición humana a varias personas, en situaciones controladas y en referencia a un número reducido de temas.

El filósofo estadounidense John Searle (n. 1932) niega que el test de Turing sea una forma válida de establecer la consciencia de una máquina. Su razonamiento se basa en el «argumento de la habitación china», con el que Searle pretende demostrar que la comprensión humana no puede reducirse a un conjunto de normas mecánicas. Searle sostiene que es posible pasar el test de Turing sin entender una sola palabra de lo que se está discutiendo.

Dicho argumento reza como sigue: en una habitación hay un hombre que no habla chino. A través de una apertura en la puerta, recibe papeles que contienen mensajes escritos con caracteres chinos. A pesar de que no sabe nada de chino, dispone de un conjunto de reglas en inglés que le permiten devolver otros papeles en chino a la persona de fuera, de forma que esta creerá que está teniendo una conversación con un nativo de este país asiático. Aunque el hombre no entiende el chino, es capaz de hacer creer a la persona de fuera que sí lo comprende. A pesar de que ha pasado el test de Turing, no posee conocimiento alguno de ese idioma. Por lo tanto, el mero uso de un conjunto de reglas mecánicas no es equivalente a la comprensión: un ordenador que pase el test de Turing no dejará de ser una máquina.

El argumento de Searle tiene varios problemas que muchos científicos y filósofos se han apresurado a apuntar. Primero, es muy difícil dar con una serie de normas mecánicas que permitan responder a preguntas que parezcan presuponer que el hablante entiende la conversación. El hecho de que ningún ordenador haya conseguido pasar el test de Turing, a pesar del gran número de equipos que lo intenta cada año, da una idea de la dificultad del problema. Resulta que es muy complicado transformar todas las posibilidades de conversación entre personas en un conjunto finito de normas: hay un gran número de detalles y sutilezas que resultan muy complicados para un ordenador. Un sistema que lograra engañar a una persona sobre su supuesta humanidad tendría que poseer un conocimiento minucioso, no solo del lenguaje, sino de la naturaleza humana, y no se limitaría a unos

cuantos papeles con normas escritas en ellos. Hofstadter argumenta que, a pesar de que la persona de dentro de la habitación puede no saber chino, el conjunto, que incluye tanto a la persona como a las normas que está utilizando, sí lo hace. De la misma forma que ninguna de nuestras neuronas sabe hablar pero juntas son capaces de generar el habla, o que la inteligencia agregada de todas las hormigas en un hormiguero es muy superior a la de cualquiera de sus miembros por separado, el conocimiento de la persona en la habitación no puede considerarse aislado. Searle comete el error de fijarse solo en la persona, sin darse cuenta de que el conocimiento puede estar distribuido en todo el sistema: es decir, Searle asume lo que quiere demostrar, lo que quita potencia a su argumento.

Ataques como el de Searle son comunes en filosofía, donde a veces se usa el término «reduccionista» de forma peyorativa. El problema con este tipo de argumentos es que, si bien pueden tener cierto éxito atacando la empresa científica de entender la consciencia, no ofrecen un modelo alternativo. Si el test de Turing no sirve para determinar la consciencia de una máquina ¿qué deberíamos usar en su lugar? Searle no se pronuncia. De hecho, muchos pensadores sostienen que no hay ninguna forma de determinar si una máquina es consciente, porque una máquina no puede serlo. El problema de este tipo de afirmaciones es que no son falsables: no hay ninguna forma de demostrar que la teoría se equivoca. Las teorías no falsables son un anatema en física, porque no son predictivas. Sin embargo, la cuestión de la consciencia es una cuestión metafísica y los estándares de la ciencia no se le aplican, por fortuna o por desgracia.

LA SUBJETIVIDAD CONTRA EL REDUCCIONISMO

El segundo argumento más citado contra el reduccionismo fue propuesto por el filósofo Thomas Nagel (n. 1937), nacido en Belgrado, hoy capital de la República de Serbia. Su idea es que la consciencia no puede ser explicada por la ciencia, porque esta es puramente objetiva, mientras que la consciencia es totalmen-

te subjetiva. Según Nagel, todo ser consciente tiene una característica en común: hay algo que denota que «se es esa persona o animal».

Nagel se pregunta si sería posible para una persona saber lo que siente un murciélago. Por supuesto, podemos imaginar que tenemos alas o que colgamos del techo. Sin embargo, tendríamos serios problemas para poder entender lo que siente el animal cuando «ve» objetos usando la ecolocalización. A pesar de que no hay duda alguna de que los murciélagos tienen sensaciones, como humanos nunca tendremos acceso a ellas. No se trata solo de que el murciélago y nosotros seamos seres distintos: estamos hablando de diferentes tipos

de sensaciones, tan distintas que no hay forma de que podamos comprenderlas a base de modificar nuestra propia percepción.

La ciencia, argumenta Nagel, es la rama del conocimiento que hace abstracción de lo subjetivo para llegar a una realidad objetiva comunicable para todos. Cuando la física describe un sonido, abstrae la sensación de oír cierto tono agudo o disonante y se centra en las características de este, que son analizables sin recurso a la subjetividad: se trata de una onda de presión, donde el aire sufre una serie de compresiones y rarefacciones, transmitiendo energía en el proceso. Pero la descripción del sonido en términos de átomos y moléculas no puede decirnos lo que se siente al escuchar una sinfonía: se limita a ignorar el problema.

Sin embargo, es incuestionable que las experiencias del murciélago y las nuestras al oír un sonido existen. Tienen una existencia obvia y, de hecho, es la única que nosotros, como humanos, podemos verificar directamente: así como la existencia de un mundo externo a nosotros podría ser una ilusión, nuestras percepciones de este no lo son. Pero ¿cómo puede la física, que se ocupa exclusivamente de lo objetivo, de lo comunicable a todo individuo, dar cuenta de la experiencia subjetiva? Parece que nuestras herramientas científicas no son las adecuadas para la tarea.

Resulta que un misterioso caos acecha tras una fachada de orden... y en cambio, dentro del caos se esconde un orden aún más misterioso.

DOUGLAS R. HOFSTADTER

Podemos imaginar, por ejemplo, una raza alienígena que visita nuestro planeta y se dedica a analizar la fisiología humana. Supongamos, también, que su tecnología es tal que les permite saber todo lo posible sobre los procesos neuronales de nuestro cerebro. ¿Les permitirá eso saber lo que siente un humano cuando llora o prueba un helado? Al fin y al cabo, la descripción de los procesos será objetiva y comunicable, mientras que la naturaleza de nuestra experiencia es justamente lo contrario.

Los argumentos de Nagel son difíciles de atacar porque hacen uso de una intuición muy potente, tanto que parece obvia. Esto es un rasgo común en todos los argumentos contra el reduccionismo: dado que estamos intentando entender el origen mismo de nuestra experiencia subjetiva, es casi imposible tratar el problema con objetividad, dado que todo lo que sabemos sobre la consciencia es por definición subjetivo.

Hofstadter neutraliza el argumento de Nagel usando el lenguaje para reinterpretar el problema y salir de lo que considera una trampa del lenguaje. Si en lugar de pensar en personas y murciélagos lo hacemos en inteligencias artificiales, la respuesta parece mucho más clara. Por ejemplo, podríamos tener dos inteligencias artificiales distintas: una tendría como sensor un sónar, mientras que la otra utilizaría una cámara de infrarrojos. Nos preguntamos ahora si la segunda máquina puede saber lo que «siente» la primera. En este caso, la pregunta se puede responder con facilidad usando el concepto de *emulación*. Una máquina emula a otra cuando realiza exactamente las mismas operaciones matemáticas. ¿Puede la primera máquina emular a la segunda? Depende. Si tiene suficiente memoria y capacidad de proceso, sí. Si no, puede que lo logre a menor velocidad o quizá sea incapaz de ello.

Sin embargo, no hay nada en este ejemplo que vaya más allá de lo físico: se trata de una pregunta específica sobre ciertas características físicas de una máquina, nada más y nada menos. De la misma forma, podemos traducir la pregunta inicial a «¿puede una mente humana emular la de un murciélago?». Probablemente no: el cerebro, al contrario que un ordenador de sobremesa, no es programable más de una vez, porque su estructura física determina su programación. Quizá se podría tomar parte de un

cerebro humano y reconectar sus neuronas para que tuviera la misma estructura que el de un murciélago, pero entonces ya no tendríamos un humano sino un quiróptero, haciendo imposible la comunicación. Pero eso no quiere decir que haya algo irreducible en la experiencia de ambos: se trata simplemente de un problema práctico, de cómo están contruidos nuestros cerebros, y no de un problema fundamental sobre la capacidad de la ciencia para dar cuenta de lo subjetivo.

Hay que recalcar que Nagel, en su artículo, no rechaza el reduccionismo: se limita a constatar que, con nuestro conocimiento actual, parece muy difícil poder explicar la emergencia de las sensaciones subjetivas, pero afirma que puede que sea posible en el futuro, aunque ignora cómo. En este sentido, Nagel se encuentra más cerca de científicos como Penrose, que aspiran a explicar la consciencia en términos de interacciones físicas, que de Searle.

LO QUE ARGUYEN LOS ZOMBIS FILOSÓFICOS

Otro argumento clásico contra el reduccionismo es el de los zombis, del filósofo analítico australiano David Chalmers (n. 1966). Su argumento se basa en la posible existencia de lo que llama «zombis filosóficos», seres que se comportan de la misma forma que una persona, pero que no tienen consciencia. Chalmers afirma que el mero hecho de que los zombis filosóficos sean concebibles implica que el reduccionismo no puede ser correcto.

Veamos por qué la posibilidad de los zombis es tan catastrófica para la visión mecanicista del mundo. Imaginemos que existe otro universo con otra Tierra poblada de gente que se comporta de la misma forma que nosotros, pero que no tiene consciencia. Donde nosotros vemos un color, los zombis actúan como si lo hubiesen visto, pero en realidad nunca conocerán la sensación de ver «rojo».

Si admitimos que un universo así es posible, también nos veremos obligados a admitir que el nuestro tiene algo extraordinario que no viene dado por las leyes de la física, ya que ambos universos son indistinguibles excepto por el hecho de que nosotros

somos conscientes. Por lo tanto, el reduccionismo no puede ser cierto y la consciencia tiene que ser algo añadido a las leyes de la física.

Por supuesto, uno puede negar de entrada que el universo de los zombis sea posible. Sin embargo, Chalmers argumenta que el mero hecho de que los zombis sean concebibles los hace posibles. Su razonamiento puede resumirse de la siguiente manera:

Los zombis filosóficos son concebibles. Cualquier cosa concebible es posible. Por tanto, los zombis filosóficos son posibles.

Es decir: no existen elefantes rosas, pero puedo imaginarlos. Eso implica que existe la posibilidad de que existan: podría imaginar un universo donde los elefantes tengan la piel rosa. Así pues, los elefantes rosas son posibles.

El argumento de Chalmers es sorprendentemente difícil de dismantelar. A pesar de que, *a priori*, los zombis pueden parecer imposibles, el filósofo propone varias situaciones en las que su existencia sería obvia. Por ejemplo, podríamos tomar el cerebro de un individuo y sustituir todas las neuronas por dispositivos controlados por un teléfono móvil. Luego distribuiríamos los teléfonos móviles entre miles de personas, a las que daríamos instrucciones para apretar un botón siempre que recibiesen determinadas señales, de forma que el cerebro funcionase de la misma forma que el de la persona a la que hemos sustituido. Podemos ahora preguntarnos si la persona realizará las mismas acciones que antes de ser reemplazada por una red de dispositivos y la respuesta será afirmativa. Sin embargo, si nos preguntamos si la persona es consciente, la respuesta no está tan clara. Incluso si adoptamos una postura radicalmente reduccionista y afirmamos que sigue siendo consciente, nos vemos obligados a admitir que el hecho de que no lo sea es concebible y, por lo tanto, posible. Así pues, debe de haber algo más que leyes de la física para explicar la consciencia.

Hay varias formas de atacar el argumento de Chalmers. La primera es rechazar que todo lo concebible es posible. Por ejemplo, se puede concebir un número par que no sea la suma de dos pri-

mos, pero eso no significa que no haya una posibilidad de que ese número exista. Chalmers sale al paso de este argumento definiendo lo concebible como todo aquello que no tiene por qué ser falso *a priori*, es decir, antes de realizar experimento alguno. Dado que las matemáticas son todas *a priori*, el argumento queda neutralizado.

Otra posibilidad es negar que los zombis sean concebibles. Esta posibilidad parece ir en contra de nuestra intuición, ya que es obvio que podemos pensar en una criatura que responde como un humano pero no tiene consciencia alguna. Sin embargo, si pensamos en ello un poco más a fondo, veremos lo difícil que es mantener nuestra posición sin contradecirnos.

Imaginemos, por ejemplo, que en el mundo zombi uno de ellos dice «no me apetece ir al cine». ¿Cómo interpretamos esa frase? En nuestro mundo, la respuesta está clara: cuando ese zombi piensa en ir al cine, puede que sienta pereza o que prefiera hacer otra cosa. Sin embargo, ni una cosa ni la otra. ¿Estará mintiendo? En ese caso, implicaría que ese ser tiene algún concepto de verdad, lo que no es posible si se trata de uno de esos «muertos vivientes». Quizá se trate de una frase sin significado, aunque tiene que referirse a algo, porque lo cierto es que nuestro zombi no va al cine.

Al final, nos vemos obligados a aceptar que su afirmación debe referirse a algún tipo de estado interno, y que en su cerebro hay una cierta configuración que corresponde a «no me apetece ir al cine». Nos podríamos preguntar también si esta persona reflexiona: ¿quizá existen los zombis filósofos? Según Chalmers, un Einstein zombi es perfectamente posible. Pero ¿cómo ha llegado a la teoría de la relatividad? Tendríamos que aceptar que una serie de procesos mecánicos le han llevado a afirmar que existe la dilatación temporal, pero que no ha habido ningún pensamiento asociado.

De la misma forma que el argumento de Searle, el razonamiento de Chalmers, aunque potente, no convence a todo el mundo y no puede considerarse una demostración de que el reduccionismo sea incorrecto.

A veces la gente que nadie imagina es la que hace las cosas que nadie puede imaginar.

ALAN TURING

Los ataques de Searle, Chalmers y Nagel al reduccionismo tienen algo en común: todos plantean críticas, pero ninguna alternativa. Aunque los tres afirman que la física no es suficiente para explicar la consciencia, ninguno propone una teoría alternativa que explique qué la causa y cómo. Tampoco tienen forma de explicar por qué los estados físicos se corresponden tan bien con los estados subjetivos de consciencia. Bueno, cabe recalcar que los tres son filósofos, no científicos. Mientras que en física gana la teoría que explica el mayor número de fenómenos con el menor número de supuestos, ese no es necesariamente el caso en otras disciplinas.

IMAGINEMOS QUE LA MENTE ES UN ORDENADOR

Los argumentos filosóficos contra el reduccionismo no son suficientemente potentes para excluir la posibilidad de que la mente sea simulable. Existen, sin embargo, argumentos matemáticos contra la posibilidad de que la mente sea un ordenador clásico. Estos fueron acuñados por el físico inglés Roger Penrose, que los publicó por primera vez en su libro *La nueva mente del emperador* (1989).

Para abordar el argumento de Penrose, es necesario explicar primero qué es un ordenador clásico y por qué su definición es relevante a la hora de explorar la mente.

En 1936, Alan Turing presentó lo que hoy se conoce como «máquina de Turing», que sentó las bases de la informática moderna. La idea de Turing era que, con una máquina capaz de realizar un conjunto reducido de operaciones básicas, tendría que ser posible realizar cualquier operación matemática. De forma algo simplificada, uno puede imaginar una máquina de Turing como un ordenador infinitamente potente y con una cantidad ilimitada de memoria. En su concepción, la máquina de Turing se basaba en un cabezal que escribía unos y ceros en una cinta infinitamente larga. Aunque parezca increíble, tanto el cabezal y la cinta como un ordenador son capaces de realizar exactamente las mismas operaciones.

Los típicos ordenadores de sobremesa realizan una cantidad minúscula de operaciones distintas. Los datos que manipulan son cadenas de unos y ceros, cada una de las cuales representa un número o un carácter. Con una cadena lo suficientemente larga de unos y ceros es posible representar cualquier número: dado que, con un número suficientemente grande, se puede codificar cualquier mensaje, las cadenas de unos y ceros sirven para representar desde textos hasta películas.

Las operaciones que un ordenador realiza en los unos y ceros son también muy limitadas. A cada uno o cero se le llama un *bit* de información; esos bits son procesados mediante dispositivos electrónicos llamados *puertas lógicas*, que toman uno o dos bits y producen un bit de salida. Las puertas lógicas se llaman así porque se basan en la lógica matemática: la idea es que cada bit nos dice si una frase es verdadera o falsa. Un uno significa «cierto» y un cero, «falso». Podemos entonces operar con estos unos y ceros para decidir si ciertas combinaciones de frases son también verdaderas o falsas.

Pongamos que tenemos una frase que sabemos que es cierta, como por ejemplo «los días tienen veinticuatro horas», y la representamos por un bit cuyo valor será uno. Si ahora creamos otra frase relacionada con la primera, del tipo «los días *no* tienen veinticuatro horas», el hecho de añadir ese «no» ha convertido una frase cierta en falsa. De la misma forma, podemos usar un «no» para transformar una frase falsa en cierta. Ese «no» añadido se denomina «operador lógico», una función que invierte el valor de verdad de una frase, como se ilustra en la figura 1. Si una frase es cierta, el operador «no» la convierte en falsa y viceversa.

FIG. 1

Entrada	Salida
0	1
1	0

El operador «no» toma como entrada un bit y lo invierte.

FIG. 2

Entrada 1	Entrada 2	Salida
0	0	0
0	1	0
1	0	0
1	1	1

Tabla de verdad para el operador «y». Este solo devuelve un resultado cierto si ambas entradas son ciertas.

FIG. 3

Entrada 1	Entrada 2	Salida
0	0	0
0	1	1
1	0	1
1	1	1

Tabla de verdad para el operador «o». Este devuelve un resultado cierto si una de las entradas lo es.

En un ordenador, esto se puede implementar mediante un circuito. El número uno se representa por un cierto valor de la corriente y el cero por otro. El circuito «no» transforma la corriente que representa un uno en un cero y viceversa.

De la misma forma, existen operadores lógicos a dos bandas, que toman dos valores de verdad como entrada y devuelven un solo valor de salida. El más típico es el operador *y*, más conocido como «AND» por su nombre inglés. Este operador toma dos valores de verdad A y B y devuelve un uno si «A y B» es cierto y un cero si no lo es. Por ejemplo, «los días tienen veinticuatro horas» y «el agua hierve a cien grados» son ciertas: por lo tanto, «los días tienen veinticuatro horas y el agua hierve a cien gra-

dos» también lo es. En cambio, «los días tienen veinticuatro horas y los cerdos son azules» es falsa: basta con que una de las dos frases sea falsa para que el valor de verdad de su combinación también lo sea, como puede verse en la tabla de verdad para el operador «y» (figura 2).

El último operador lógico que necesitamos es el operador «o» u «OR» en inglés (figura 3). Este funciona como el operador «y», pero devuelve un valor cierto si cualquiera de sus dos frases de entrada es cierta. Por ejemplo, «los días tienen veinticuatro horas o el agua hierve a cincuenta grados» es cierta, a pesar de que «el agua hierve a cincuenta grados» no lo es.

¿UNOS Y CEROS? NO PARA TODO

Resulta increíble que, con solo tres operadores lógicos, sea posible realizar cualquier operación matemática. De hecho, es posible hacerlo con menos, pero explicarlo queda fuera del ámbito de este libro. Dado que las operaciones matemáticas son todo lo que necesitamos para crear sistemas de complejidad arbitraria, hemos encontrado una forma de crear una máquina universal, que puede procesar cualquier tipo de información y transformarla en lo que queramos. Esta es la base de nuestros ordenadores personales, que hoy damos por sentados, pero que no existirían de no ser por la idea de Turing.

Dado que las leyes de la física son matemáticas, tendría que ser posible introducir el universo entero en una máquina de Turing: los únicos problemas posibles son el espacio y el poder de procesamiento pero, dado que la máquina de Turing es un ordenador idealizado, tenemos una cantidad infinita de ambos. Así pues, en principio podríamos determinar el estado del universo en un cierto instante, guardarlo en nuestra máquina de Turing, programarla con nuestras leyes de la física, que no son más que normas para la manipulación de unos y ceros, y ejecutar nuestro programa. Si el universo está basado en leyes matemáticas, tiene que ser simulable, porque una máquina de Turing es capaz de realizar cualquier operación matemática. Por supuesto, el uni-

OPERACIONES MATEMÁTICAS CON PUERTAS LÓGICAS

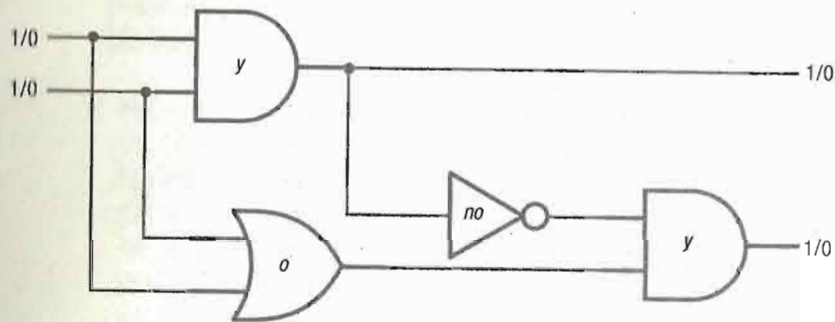
Las puertas lógicas son dispositivos electrónicos concebidos para realizar operaciones lógicas, como sumar o multiplicar, negar o afirmar, incluir o excluir. Se basan en la denominada álgebra de Boole o booleana, una estructura algebraica esquematizada en «Y», «O», «NO» y «SI». Estas compuertas ofrecen la posibilidad de efectuar cualquier operación matemática con ellas. Un primer paso es poder multiplicar o sumar dos números binarios, de un bit cada uno. Cuando esas dos operaciones se pueden realizar, el resto es solo una cuestión de añadir complejidad al circuito.

¿Lo más fácil? Multiplicar

Sorprendentemente, la multiplicación es la más sencilla de las operaciones: basta con una puerta y. Si tomamos dos números que pueden ser uno o cero, obtendremos cero siempre que uno de los dos sea cero; y uno si ambos son uno. Esto es lo mismo que sucede con la tabla de verdad de la puerta y: solo obtendremos «verdadero», el número uno, si ambas frases son ciertas, y falso para el resto.

Cuando las sumas se embrollan

Para la suma se requiere un circuito un poco más complicado, que se muestra en la figura. Obtener el primer bit, en la parte superior derecha, es sencillo: uno más uno es dos, que se expresa como «10» en binario. Así que el primer bit será uno si los dos bits de entrada lo son. En cambio, calcular el segundo bit es mucho más difícil y requiere la combinación de varias puertas lógicas para obtener un valor que será uno si los dos bits de entrada son distintos.



Circuito lógico para sumar dos bits. La puerta y superior da el primer dígito, que será 1 si ambos números son 1. La combinación inferior de puertas es necesaria para obtener 1 cuando los dos bits son distintos.

verso nos incluye, lo que significa que la mente humana también tiene que poder ser reducida a un conjunto de operaciones entre unos y ceros.

Penrose argumenta, sin embargo, que los humanos son capaces de realizar operaciones que son imposibles para una máquina de Turing, lo que significa que no es posible reducir nuestro comportamiento a unos y ceros. Dado que la máquina de Turing puede ejecutar cualquier operación matemática, eso significa también que nuestra consciencia no puede reducirse a la aplicación de una serie de leyes matemáticas.

EL TEOREMA DE COMPLETITUD DE GÖDEL

Uno de los teoremas matemáticos más influyentes del siglo xx fue el *teorema de Gödel*, planteado por el matemático y filósofo Kurt Gödel (1906-1978), de origen austrohúngaro y nacionalizado estadounidense, en el que se basa el argumento de Roger Penrose. Para entender lo que dice, debemos primero hablar de algunos principios fundamentales de la matemática.

A pesar de que las matemáticas que se suelen aprender en la escuela versan sobre números, en realidad estas constituyen un *sistema formal*: un conjunto de símbolos y de reglas para manipularlos. La idea es que hacer matemáticas es como jugar al ajedrez: se tienen unas cuantas piezas, cada una de las cuales sigue un conjunto de reglas. Estas reglas podrían ser otras: las elegimos simplemente porque son las que hacen el juego más interesante. En este sentido, el hecho de que las matemáticas se correspondan con la realidad es irrelevante: podemos tener teorías matemáticas que hablen de entidades que nunca veremos en este universo y estas no dejarán de ser ciertas, en el sentido matemático del término. Es deseable que las matemáticas se correspondan con la realidad, pero no es esencial. Según la visión formalista de las matemáticas, una teoría matemática está constituida por tres elementos básicos: *símbolos*, *axiomas* y *reglas de inferencia*. Los símbolos matemáticos son como las palabras de la teoría: los podemos unir para construir afirmaciones

que serán ciertas o falsas. Si son ciertas, las llamaremos *teoremas*. Los axiomas son las afirmaciones iniciales que suponemos ciertas y a partir de las cuales derivaremos el resto de afirmaciones de la teoría. Por ejemplo, la geometría euclídea cuenta con cinco axiomas, el más famoso de los cuales afirma que las paralelas nunca se cortan; es posible construir geometrías alternativas sustituyendo este axioma por otro. De hecho, parece que se ajustan mejor a la realidad que la euclídea.

Finalmente, tenemos las reglas de inferencia, que nos permiten derivar nuevas afirmaciones a partir de los axiomas. Combinando axiomas y reglas de inferencia podemos obtener teoremas, que se pueden usar a la vez para obtener nuevos teoremas y así sucesivamente.

Volviendo a la metáfora del ajedrez, las piezas serían los diferentes símbolos, mientras que las reglas del juego serían las reglas de inferencia. La posición inicial de las piezas es el axioma, mientras que todas las configuraciones del tablero a las que se puede llegar aplicando las reglas son teoremas.

Hay ciertas propiedades de una teoría matemática que son esenciales: la primera es que sea «consistente», es decir, que no sea posible demostrar una cosa y su contraria. Hay un teorema de la lógica que afirma que, de una contradicción, se deriva cualquier cosa: es decir, si nuestra teoría contiene una contradicción, puede usarse para demostrar lo que queramos, lo que le quita cualquier tipo de poder predictivo. Así pues, una buena teoría matemática tiene que ser consistente. Sin embargo, demostrarlo es sorprendentemente difícil, así que en muchos casos los matemáticos se limitan a asumir que una teoría lo es hasta que se demuestre lo contrario.

Otra propiedad importante para una teoría matemática es su *completitud*: la idea es que, si existe una afirmación que es cierta en la teoría, debería ser posible demostrarla usando solo sus axiomas y reglas de inferencia, es decir, sin utilizar nada externo al sistema. Si no, nos veríamos con un problema inmenso: para demostrar ciertas proposiciones en nuestra teoría, necesitaríamos axiomas extra que habría que introducir a mano, pero esos axiomas extra introducirían nuevas proposiciones indemostra-

bles, para las cuales necesitaríamos nuevos axiomas, y así hasta el infinito.

A principios del siglo xx, entre los matemáticos existía la creencia extendida de que toda teoría matemática debía ser consistente y completa. Sin embargo, Kurt Gödel demostró que eso era imposible. Una teoría matemática puede ser consistente o completa, pero no ambas cosas a la vez.

El teorema de Gödel posee una gran complejidad matemática, así que aquí se dará solo una explicación intuitiva de en qué se basa. Para ello, usaremos la idea de una máquina de Turing como modelo para una teoría matemática.

Como se ha visto, una teoría matemática no es más que una serie de símbolos, axiomas y reglas de inferencia.

Dado que los tres pueden expresarse en el lenguaje, pueden ser traducidos a una combinación de unos y ceros. Así pues, si uno introduce todos los ingredientes para la teoría en un ordenador con capacidad y potencia infinitas, tendrá algo equivalente a la teoría matemática que está desarrollando. Es decir: el ordenador tendría que ser capaz, a base de fuerza bruta, de realizar las mismas tareas que un ejército de matemáticos competentes.

Ahora podemos usar el ordenador, con la teoría cargada en la memoria, para averiguar si ciertas afirmaciones son teoremas. Si, por ejemplo, introdujéramos la geometría euclídea y preguntásemos si los ángulos en un triángulo suman 180 grados, nos diría que la afirmación es correcta. En caso de que nuestra teoría sea consistente, el ordenador no se contradecirá nunca; si es completa, el ordenador siempre será capaz de responder afirmativamente o negativamente a cualquier cuestión sobre nuestra teoría.

Gödel tuvo la idea de que, si una teoría es lo suficientemente potente como para poder referirse a sí misma, incurrirá en ciertas paradojas que crearán inconsistencias en ella. Una forma intuitiva de presentar su idea es imaginar que introducimos en el ordenador una afirmación que se refiere a sí misma, a la que lla-

Todos los principios de la matemática se reducen a principios de la lógica.

WILLARD VAN ORMAN QUINE
EN *DESDE UN PUNTO DE VISTA LÓGICO*

maremos G, y a la que le preguntaremos si es verdadera. La frase G dice: «El ordenador nunca dirá que G es cierta» o, en otras palabras, «el ordenador nunca dirá que esta frase es cierta».

Veamos cómo contestará el ordenador. Si dice que G es cierta, entonces G es falsa, porque el ordenador ha dicho que es cierta. Por lo tanto, no puede decir que es cierta, pero entonces G es verdadera, porque el ordenador nunca lo ha dicho. En otras palabras: tenemos una afirmación, G, que es cierta, pero nuestro sistema no puede decírnoslo. Por lo tanto, nuestro sistema es incompleto. Alternativamente, podríamos dejar que el ordenador se pronunciase, pero entonces se contradecirá a sí mismo, mostrándose inconsistente.

Nuestro sistema puede ser o completo o consistente, nunca las dos cosas a la vez. Viendo la explicación anterior podría parecer que hemos hecho trampas, porque es así. Al implicar al propio ordenador en la pregunta, hemos creado una paradoja. Sin embargo, una teoría no puede referirse a sí misma, o eso puede parecer en un principio. El genio de Gödel consistió en encontrar una forma de lograr que una teoría se refiriese a ella misma: en particular, Gödel tradujo uno a uno los símbolos de la teoría y una serie de números enteros para que la teoría pudiese referirse a sí misma. A ese proceso se le llama hoy *gödelización*. Por eso, el teorema de Gödel solo se aplica a teorías matemáticas lo suficientemente potentes como para contener la aritmética: es necesaria esa estructura para poder realizar la traducción «de fuera a dentro». Es decir, para poder escribir afirmaciones sobre el propio sistema desde dentro del sistema mismo.

Podemos intentar entender la situación retomando la metáfora del ajedrez. Si introducimos todas las reglas del juego en un ordenador, podríamos esperar que este llegue a todas las configuraciones posibles del tablero, lo que parece dar cuenta de todos los teoremas posibles. Sin embargo, las personas somos capaces de ir más allá: por ejemplo, en ciertas configuraciones (figura 4), somos capaces de deducir que el alfil nunca pisará una casilla negra. Se trata de una regla general sobre ajedrez que podemos extraer de casos particulares: si consideramos toda configuración como un teorema, nuestra afirma-

LA GÖDELIZACIÓN

Para hacer que su sistema pudiese referirse a sí mismo, Gödel encontró una manera de codificar cualquier cadena de símbolos en un número natural, de forma que todo teorema de la aritmética tuviese una expresión dentro de la aritmética. Primero, Gödel asignó un número natural a cada símbolo de la teoría, para poder identificarlos unívocamente. La siguiente parte era encontrar una forma de guardar el orden de los símbolos en un teorema dentro del número natural. Para ello, tuvo la idea de utilizar los números primos, dado que estos siguen un orden determinado. Pongamos que queremos almacenar la siguiente expresión en un número:

$$x > 0 \rightarrow 0 < x.$$

Esto se lee como «x es mayor que cero implica que cero es menor que x». Podemos ahora asignar un número arbitrario a cada símbolo, como muestra la tabla. Así podremos expresar nuestra cadena de símbolos como un número, usando los números primos para almacenar la posición. Cada primo se eleva al número correspondiente al símbolo que se encuentra en esa posición en la expresión.

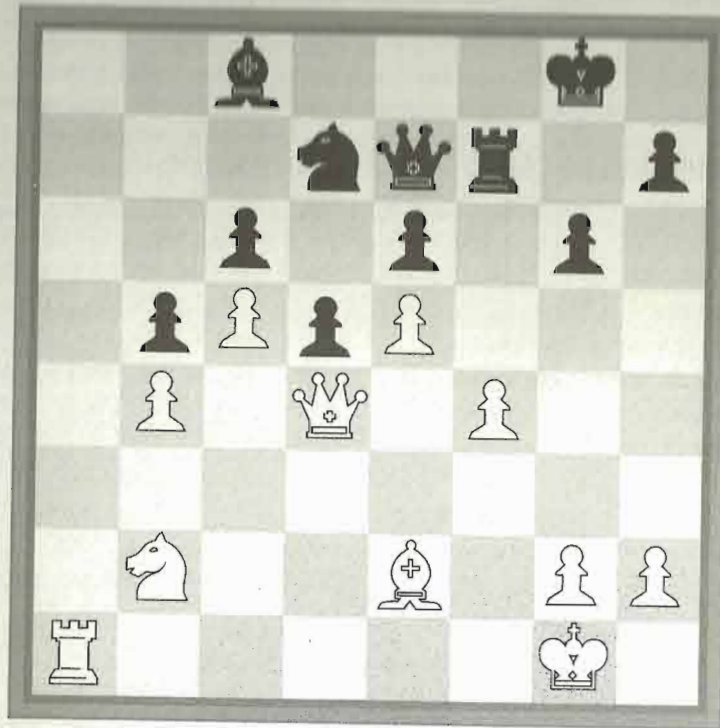
$$\text{Número} = 2^1 3^3 5^0 7^1 11^2 13^0 17^1 = 4,538563504 \cdot 10^{16}.$$

La cadena original de símbolos se puede reconstruir a base de descomponer nuestro número en factores primos: hemos encontrado un número que representa de forma inequívoca nuestro mensaje. De este modo, podemos proceder a efectuar una traducción una a una de cualquier afirmación sobre la aritmética a una relación sobre los números mismos.

Símbolo	Número
x	1
>	2
0	3
→	4
<	5

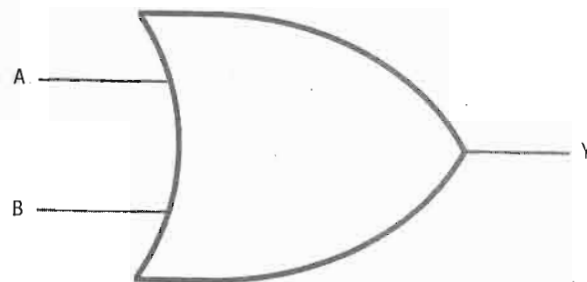
Correspondencia entre los símbolos de la expresión y sus números. Estos se pueden escoger arbitrariamente, mientras haya un único número para cada símbolo.

FIG. 4



El alfil blanco nunca pisará la casilla negra. Se trata de una norma general que no se puede deducir por fuerza bruta.

FIG. 5



Esquema de una puerta lógica. Esta toma las entradas A y B y produce una salida Y.

ción es un *metateorema*: un teorema sobre las propiedades de cada teorema. Esta habilidad de salir fuera del sistema es algo que un ordenador no tiene, si se limita a aplicar ciegamente las reglas de inferencia.

UN UNIVERSO NO COMPUTABLE PERO DETERMINISTA

Penrose cree que el teorema de Gödel muestra que los humanos somos capaces de realizar razonamientos que un ordenador clásico nunca será capaz de hacer. Nosotros podemos ver que esa frase que se autoafirma, a la que hemos llamado G, es cierta, a pesar de que el ordenador nunca nos lo dirá. De la misma forma, podemos hacer generalizaciones sobre ajedrez que el ordenador tampoco puede realizar si se limita a aplicar las reglas de inferencia. Según Penrose, eso demuestra que nuestra mente no puede ser descrita por un ordenador clásico y, por tanto, no puede ser simulada. Tiene que haber algo diferente en nuestro cerebro, algo más allá de los unos y ceros y de esas puertas lógicas que, basadas en uno o dos bits producen un bit de salida (figura 5).

Es importante recalcar que Penrose no rechaza la idea de que la mente sea resultado de procesos físicos: lo que rechaza es que sea un producto de procesos físicos clásicos, es decir, procesos que no incluyan la mecánica cuántica. De hecho, su hipótesis es que, para explicar la mente, es necesario recurrir a fenómenos cuánticos que podrían explicar ese tipo de inferencias casi mágicas que el cerebro humano es capaz de realizar.

El argumento de Penrose puede resumirse como sigue: primero, el teorema de Gödel nos muestra que un ordenador infinitamente potente no es capaz de averiguar si G es cierta. Esto significa que el problema no se resuelve cambiando de máquina, porque la máquina de Turing es capaz de realizar cualquier operación matemática en un tiempo infinitamente pequeño. En caso de que la máquina de Turing no sea capaz, tampoco lo es ningún otro ordenador. Así que, si una persona es capaz de averiguar que la frase G es cierta, lo tiene que estar haciendo de forma no computacional. En otras palabras, la verdad o false-

dad de G es «no computable»: no existe máquina alguna que sea capaz de computar una respuesta.

Las personas somos capaces de ver que G es cierta, lo que significa que estamos realizando una operación no computable. Parece pues que tiene que haber algo no computable sucediendo en nuestro cerebro, de lo que se desprende que las personas no podemos ser simuladas por un ordenador que opere según las leyes de la física clásica.

A pesar de su visión de la mente, Penrose es un reduccionista: sigue creyendo que las habilidades de la mente tienen que ser explicables por las leyes de la física. Tanto la relatividad especial como la mecánica newtoniana son computables, así que el único candidato posible a contener procesos físicos no computables es la mecánica cuántica.

Para Penrose, el colapso de la función de onda tiene que ser ese proceso. Su tesis es que, al realizar una medida cuántica, se da un proceso determinista pero no computable, que percibimos como aleatorio porque no lo podemos predecir con un ordenador. Al fin y al cabo, uno siempre puede poner un generador de números aleatorios en un ordenador, así que si la realidad fuese puro azar, no habría ningún problema en simularla con un ordenador. El hecho de que las personas realicen operaciones no computables es una señal de que existen procesos físicos que tampoco lo son.

En este momento uno puede preguntarse cómo algo puede ser, a la vez, determinista y no computable: si está determinado, tendría que ser predecible en principio. Sin embargo, Penrose ha creado varios sistemas que tienen ambas propiedades: su evolución está determinada, pero es imposible computarla. A pesar de que esos sistemas son poco realistas, sirven para mostrar la posibilidad de un universo no computable pero determinista.

REDES NEURONALES HUMANAS

Otra de las ideas de Penrose es que, durante el colapso de la función de onda, tiene que darse un proceso determinista pero

EL PROBLEMA DE LA PARADA Y EL TEOREMA DE GÖDEL

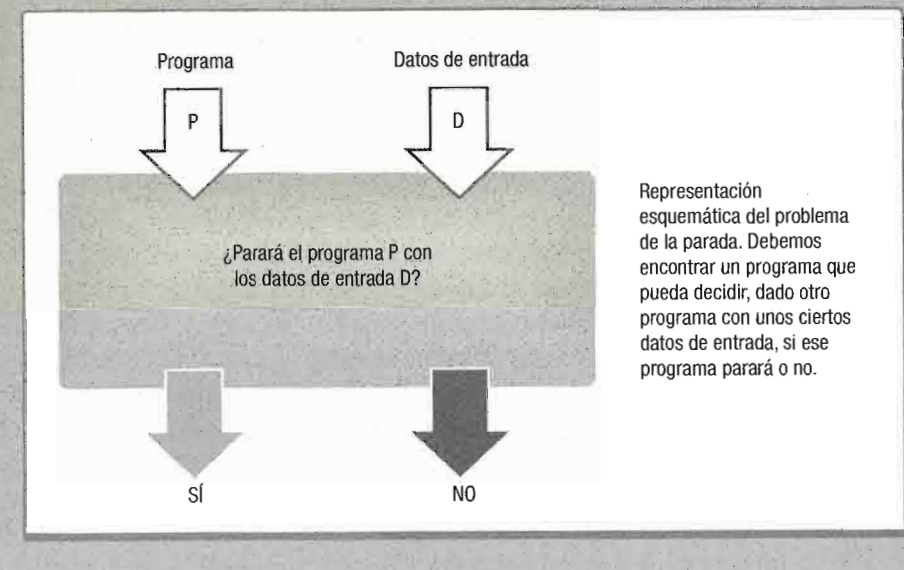
Turing reformuló el teorema de Gödel en términos de su máquina y de algo llamado el «problema de la parada», esquematizado en la figura. Este consiste en preguntarse si es posible averiguar de forma sistemática si un programa de ordenador se detendrá. Turing logró demostrar que no hay ningún procedimiento sistemático que nos lo pueda decir.

¿Una cuestión irresoluble?

Resolver el problema de la parada es equivalente a violar el teorema de Gödel. Todo teorema matemático se puede expresar como un problema de la parada. Por ejemplo, el teorema que nos dice que hay un número infinito de números primos puede enunciarse de la siguiente manera:

«El programa que se detiene cuando ha acabado de computar números primos no parará nunca».

Si hubiera una forma de resolver sistemáticamente el problema de la parada, tendríamos también una forma sistemática de demostrar cualquier teorema, ya que estos se pueden expresar como un problema de la parada. Dado que esto no es posible, como mostró Gödel, el problema de la parada no puede resolverse. La demostración de que el problema de la parada no es computable es similar a la del teorema de Gödel: Turing ideó una forma para que el algoritmo que decide si otro algoritmo parará se refiriese a sí mismo, creando una paradoja similar a la de su contemporáneo.



no computable: es decir, la mecánica cuántica no es realmente aleatoria, lo que sucede es que no somos capaces de calcular el siguiente estado aunque el universo sí que puede. En su opinión, ni la interpretación de Copenhague ni la de los múltiples universos están en lo cierto. Coincide con la de Copenhague en que el colapso de la función de onda es real, pero afirma que este se produce debido a la influencia de la gravedad a escalas microscópicas. Es decir, los efectos gravitatorios causan una inestabilidad en la función de onda que lleva a su colapso; ese proceso es determinista pero imposible de calcular.

La teoría de Penrose empezó meramente como un ejercicio teórico basado en el teorema de Gödel, pero sus afirmaciones sobre la naturaleza cuántica de la consciencia necesitaban una base experimental más sólida para ser tomadas en serio. La mayoría de físicos y biólogos están convencidos de que cualquier proceso cuántico en el cerebro desaparecerá casi instantáneamente debido a la decoherencia: el cerebro humano es demasiado grande, caliente y húmedo como para que puedan darse fenómenos cuánticos significantes en él.

A mediados de los noventa Penrose empezó a trabajar con el anestesista Stuart Hameroff (n. 1947), según el cual los microtúbulos, unas pequeñas estructuras tubulares que se encuentran en las células eucariotas (aquellas que presentan un núcleo diferenciado mediante una membrana), podrían ser los responsables de la consciencia. Hameroff sostiene que la anestesia funciona porque ataca directamente esas estructuras, suficientemente complejas como para realizar computaciones. De ello deduce que son precisamente los microtúbulos los que producen la consciencia, en lugar de propiciar meras interacciones entre neuronas, como sostienen otros neurólogos. Tanto Penrose como Hameroff están convencidos de que el interior de los microtúbulos está lo suficientemente aislado de su entorno como para evitar una decoherencia instantánea.

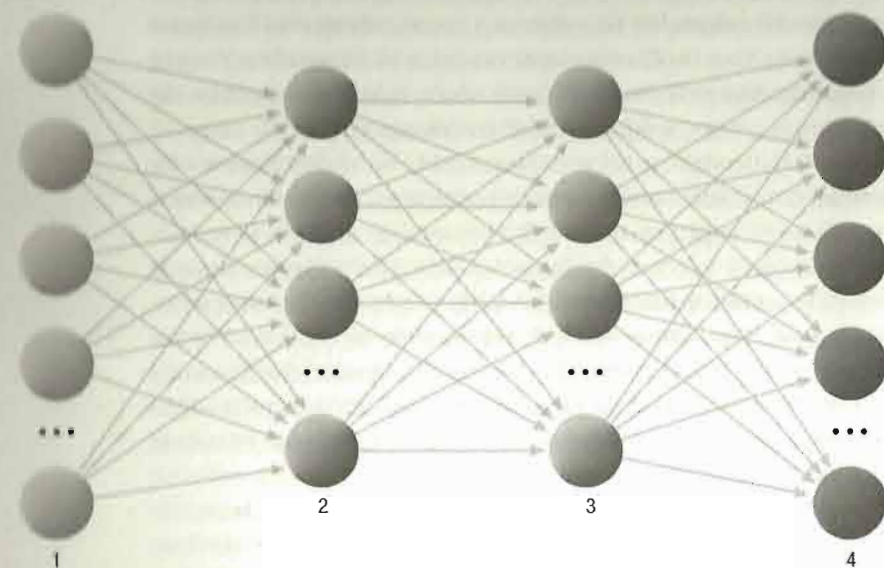
La hipótesis de Penrose y Hameroff ha recibido numerosas críticas. El cosmólogo Max Tegmark (n. 1967), un famoso defensor de la hipótesis del multiverso, término que define los hipotéticos universos múltiples existentes, ha calculado que la decoherencia

tendría que producirse en 10^{-18} segundos, lo que haría imposible que los procesos cuánticos en los microtúbulos tuvieran efecto alguno. De momento, no hay ninguna prueba de que la hipótesis de Penrose y Hameroff sea cierta, aunque sí se ha descubierto que los procesos cuánticos en seres vivos son más comunes de lo que se creía, por ejemplo en la fotosíntesis.

Además de las críticas experimentales, la teoría de Penrose ha sido atacada por sus argumentos matemáticos. En particular, una parte clave del argumento de Penrose es que un ordenador no puede afirmar la verdad de cierto tipo de afirmaciones porque no puede contradecirse a sí mismo. Sin embargo, los seres humanos parecen ser capaces de creer en todo tipo de contradicciones, sin que eso les represente problema alguno. ¿Significa eso que no podemos ser simulados por un ordenador? ¿O es posible desarrollar sistemas informáticos que se contradigan de la misma forma que lo hacemos nosotros?

La respuesta a la segunda pregunta parece estar bastante clara: las personas no se comportan como un ordenador clásico, sino que son mucho más parecidas a una *red neuronal*. Las redes neuronales son dispositivos inspirados por el cerebro humano capaces de realizar muchas de las tareas que hasta la fecha se consideraban exclusivas de las personas. Una red neuronal (figura 6) se puede modelar como un circuito de varias capas, donde la primera recibe una serie de estímulos, que procesa y envía a la segunda capa y así sucesivamente. La idea es que, a medida que avanzamos por la red neuronal, la información se agrupa en categorías más y más abstractas.

Las redes neuronales están basadas en lo que sabemos del cerebro, en particular sobre el proceso de imágenes. Una persona es capaz de reconocer a un amigo entre la gente, aunque ese amigo esté mal enfocado o en la sombra. De la misma forma, podemos distinguir entre una mesa y un taburete, a pesar de que hay mesas que parecen taburetes sobredimensionados. Esta capacidad para lidiar con la ambigüedad resultaba un misterio hasta hace relativamente poco; hoy en día, los sistemas de inteligencia artificial basados en redes neuronales se acercan cada vez más a los humanos y, en ciertas tareas, los superan.



Una red neuronal típica. La primera capa recibe los estímulos primarios, que procesa y envía a la segunda capa. Cada capa realiza un proceso añadido, hasta que la última capa emite una señal de salida. Cuantas más capas, más efectiva es la red neuronal.

El psicólogo experimental canadiense Steven Pinker (n. 1954) explica en su libro *Cómo funciona la mente* el funcionamiento de la red neuronal asociada a nuestra visión. Nuestro sistema visual tiene un gran número de capas, cada una asociada a ciertas características de la imagen. Por ejemplo, la primera capa está configurada para detectar líneas a ciertos ángulos: una determinada neurona solo se activará si ha detectado una línea con una inclinación determinada. La segunda capa no tomará los puntos de luz como entrada, sino esas líneas que han procesado las neuronas de la primera capa. Usando estas líneas, las neuronas se activan cuando ven ciertas formas. De nuevo, la salida de esas neuronas va a otra capa, cuya salida es algo aún más abstracto.

Usando esas capas de abstracción progresiva, las personas somos capaces de identificar multitud de objetos. Sin embargo, esa

potencia tiene un precio: cometemos errores. Una red neuronal no está hecha para ser infalible, sino para que sea capaz de acertar un número razonable de veces. Por eso a veces confundimos a una persona con otra o creemos ver un animal donde solo hay una sombra. Nuestra red neuronal no está hecha a prueba de fallos, porque su cometido no es la precisión matemática, sino la efectividad. Nuestros cerebros tienen que tomar decisiones instantáneas con información fragmentada, aunque se equivoquen a veces. Uno siempre está a tiempo de cambiar de opinión.

La forma de funcionar de nuestras redes neuronales sugiere que el modelo de la máquina de Turing no se aplica a la mente humana: más bien somos un procesador estadístico que hace inferencias probables a partir de información incompleta. Por consiguiente, no es de extrañar que seamos capaces de poseer intuiciones que parecen casi mágicas para un ordenador: procesamos la información de forma muy diferente, yendo de lo específico a lo general y detectando patrones recurrentes en la información que nos llega. Además, estamos exentos de requerimientos como la consistencia lógica, que es uno de los pilares de la demostración de Gödel.

¿Significa esto que las personas no podemos ser simuladas por un ordenador clásico? De momento, la respuesta parece ser que no. De hecho, las redes neuronales se ejecutan regularmente en ordenadores de sobremesa. Todo el servicio de reconocimiento de voz de Google se ejecuta en sus servidores, usando ordenadores potentes, sí, pero de un funcionamiento igual, en principio, al de una máquina de Turing. Es decir, aunque parezca que los procesos del cerebro humano sean completamente distintos a los de un ordenador convencional, la potencia de esas máquinas es tal que pueden adaptarse para simular cualquier otro tipo de lógica, aunque sea distinta de aquella con la que se han construido sus circuitos. Esto nos da una idea del poder que posee una máquina de Turing: es capaz de realizar cualquier proceso computable, incluidos aquellos que implican una forma completamente distinta de analizar información. Si lo pensamos con detenimiento, los humanos ni siquiera usamos la lógica matemática con la que se programan los ordenadores: más que en afirmaciones ciertas o falsas,

EL UNIVERSO DE POLIMINÓS

Quizá el ejemplo más sencillo de universo no computable es el «universo de poliminós». Un poliminó es un objeto geométrico que se obtiene al unir varias celdas iguales de modo que cada par de celdas vecinas compartan un mismo lado. Uno empieza con un pequeño cuadrado y va añadiendo más, probando todas las formas diferentes en las que los puede juntar. Dos poliminós que pueden transformarse el uno en el otro mediante una rotación se consideran el mismo.

El poliminó según Penrose

Hay un número infinito de poliminós, que se obtienen a base de añadir cuadrados y buscar todas las configuraciones posibles que no sean equivalentes. A pesar de que son infinitos, los poliminós pueden ordenarse. Hay muchísimas formas de hacerlo, pero la que Roger Penrose escogió se basa en el número total de cuadrados. El físico y matemático asignó un número a cada grupo de poliminós con una cierta cantidad de cuadrados, planteado en la figura 1, buscando todas las posibles combinaciones con ese número antes de llegar a la siguiente, como muestra la figura 2, con piezas de cinco.

Un problema no computable

Los poliminós son algo más que una curiosidad matemática. Con un grupo dado de ellos, ¿podríamos cubrir todo el suelo, sin dejar espacio alguno? La respuesta es trivial en algunos casos, como el de un solo cuadrado. Sin embargo, en general la solución de este problema no es computable: dado un grupo de poliminós, no hay ninguna forma sistemática de averiguar si cubren o no el plano, a pesar de que la respuesta está determinada de antemano. Penrose propone el siguiente modelo de universo determinista pero no computable: a cada instante, el estado del universo viene dado por un par de grupos de poliminós: por ejemplo, (S_n, S_0) . La evolución al siguiente instante viene dada por la siguiente norma: dado un estado (S_n, S_0) , el siguiente estado será:

- (S_{n+1}, S_0) si es posible cubrir el plano con los poliminós anteriores.
- (S_n, S_{0+1}) si no es posible cubrir el plano con los poliminós anteriores.

Como puede verse, la evolución del universo viene dada por la respuesta a una pregunta no computable, lo que significa que no es computable. Sin embargo, si es determinista: dado un cierto conjunto de poliminós, o bien cubren el suelo sin agujeros o bien no lo hacen, con lo cual todo estado futuro está unívocamente determinado. Por supuesto, nuestro universo no tiene nada que ver con el sugerido por Penrose, pero este modelo de juguete sí que hace patente la posibilidad de que algo esté determinado pero a la vez no puede ser calculado.

Una idea de Salomon Wolf Golomb

El matemático estadounidense Salomon W. Golomb (n. 1932), especializado en la resolución de problemas de análisis combinatorio y en teoría de números y de códigos, es célebre por

sus ingeniosos juegos matemáticos. Los poliminós los ideó en 1953 y sus piezas de cinco segmentos, los pentominós, fueron la inspiración del videojuego conocido como *Tetris*. El famoso rompecabezas virtual, diseñado y programado por el ruso Alekséi Pájitnov y lanzado en 1984, está compuesto por tetraminós, piezas de cuatro (*tetra* en griego) segmentos. Aunque el poliminó y el *Tetris* son parecidos, el primero puede llegar a tener mucha más complejidad.

FIG. 1

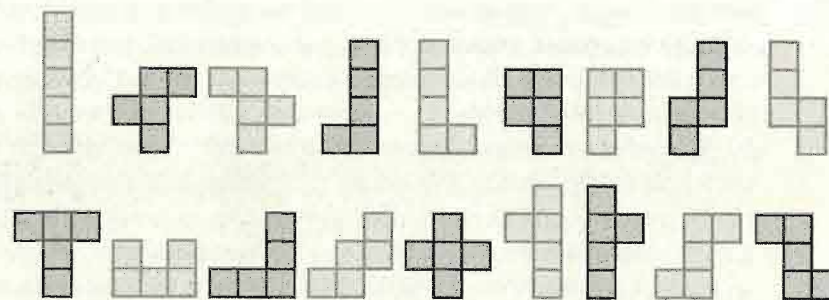
$$S_0 = \{ \square \}, S_1 = \{ \begin{smallmatrix} \square \\ \square \end{smallmatrix} \}, S_2 = \{ \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \}, S_3 = \{ \begin{smallmatrix} \square & \square \\ \square & \square \\ \square & \square \end{smallmatrix} \},$$

$$S_4 = \{ \begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \end{smallmatrix} \}, S_5 = \{ \begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \end{smallmatrix}, \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \}, S_6 = \{ \begin{smallmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \end{smallmatrix} \}, S_7 = \{ \begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \end{smallmatrix}, \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \}, \dots$$

$$S_{896} = \{ \begin{smallmatrix} \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square \end{smallmatrix} \}, \dots, S_{57984} = \{ \begin{smallmatrix} \square & \square & \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square & \square & \square \end{smallmatrix}, \dots \}$$

Los poliminós de Penrose, con su orden.

FIG. 2



Poliminós formados por cinco cuadrados, denominados «pentominós». Hay 12 distintos y 6 especulares.

nuestros estados mentales se basan en creencias, en afirmaciones que consideramos que tienen una alta probabilidad de ser ciertas. Nuestro proceso lógico está basado en esa probabilidad, mientras que el de un ordenador clásico se basa en la certeza. Sin embargo, es posible emular nuestra forma de operar usando solo operaciones entre unos y ceros y tres tipos de puertas lógicas.

Al final, la verdad o falsedad de la teoría de Penrose solo podrá ser determinada mediante experimentos. De momento, parece que no hay una cantidad suficiente de pruebas a su favor como para demostrar fehacientemente que la mente es un fenómeno cuántico; el gran éxito que están teniendo las redes neuronales en la actualidad da aún más munición a sus críticos.

UN UNIVERSO COMPUTABLE

Incluso si Penrose no está en lo cierto y la mente no es un fenómeno cuántico, hay otras posibilidades que harían imposible la simulación de la mente por un ordenador. En particular, si el universo contuviera una cantidad infinita de información, no sería posible reproducirlo a base de operaciones con unos y ceros.

En este caso, la física clásica resulta ser mucho más problemática que la cuántica. Tanto la mecánica newtoniana como la relatividad describen lo que matemáticamente se llaman «espacios continuos»: espacios en los que cualquier longitud se puede dividir en pedazos menores. Esto, que parece una necesidad obvia, causa un sinfín de problemas cuando se intenta analizar en términos de información.

Fijémonos en un simple intervalo de longitud, por ejemplo de cero a un metro. Cada posición en el intervalo viene dada por un cierto número: algunos de ellos tienen una expresión decimal que, o bien acaba en algún punto, o bien es periódica, es decir, se repite infinitamente. A este tipo de números los llamamos *racionales*. Otras vendrán dados por números cuya expresión decimal es infinita pero aleatoria: es imposible predecir el siguiente decimal a partir de los anteriores. A estos se les llama *números irracionales*.

Afirma un teorema matemático que, si tomamos un número al azar, hay una probabilidad de casi el cien por cien de obtener un número irracional. La razón para esto es que, dados dos números racionales cualesquiera, por muy cerca que estén, se puede siempre encontrar una cantidad infinita de números irracionales entre ellos. Por ejemplo, tomemos los números 5,002 y 5,003: si deseamos encontrar un número irracional, solo tenemos que escoger que empiece con 5,002 y añadir un número infinito de decimales aleatorios: tendremos un número irracional mayor que 5,002 y menor que 5,003. Esto se puede hacer para cualquier par de números racionales, dado que sus expresiones decimales paran en algún momento. En el caso de los números periódicos, tarde o temprano diferirán en al menos un decimal, tras lo cual se puede aplicar el mismo proceso que arriba.

Si el universo viene descrito por un espacio infinitamente divisible, una posición cualquiera vendrá dada por un número irracional. Pero esto es un problema porque, para almacenar un número irracional, necesitamos una cantidad infinita de información.

Supongamos, por ejemplo, que queremos almacenar el número 235 en un ordenador. Lo primero que debemos hacer es transformarlo a notación binaria, ya que los ordenadores solo entienden unos y ceros. Para simplificar, usaremos el hecho de que cada decimal ocupa aproximadamente tres bits, es decir, tres números que pueden ser o cero o uno, y calcularemos cuánta información necesitamos para almacenarlo. Dado que tenemos tres cifras y cada cifra ocupa unos tres bits, el número 235 ocupará unos nueve bits en memoria.

En cambio, la raíz cuadrada de dos ocupará mucho más. Su expresión es la siguiente:

1,41421356237...

Tenemos un número infinito de cifras y cada cifra ocupa tres bits, por lo que aunque llenáramos todos los discos duros del

planeta aún no tendríamos suficiente. Por suerte para los físicos y para los ordenadores, en realidad nunca se usan todos los decimales de un número. De hecho, cualquier medida física tiene una precisión limitada, lo que trunca cualquier número a una cantidad manejable de cifras. Por ejemplo, la carga del electrón, que es uno de los números que se conocen con más precisión, es conocida hasta los doce decimales. En general, las cantidades que usamos en las simulaciones tienen muchos menos, de forma que la estructura matemática de nuestro espacio no se manifiesta. A la práctica, el universo o, al menos, los procesos que nos suelen interesar en él, son perfectamente simulables.

Sin embargo, el hecho de que tanto la mecánica newtoniana como la relatividad contemplen espacios infinitamente divisibles sí es relevante para reflexionar sobre si, en principio, el universo puede ser computado o si, por el contrario, cualquier parte, por pequeña que sea, contiene una cantidad infinita de información. El espacio clásico nos indica que la respuesta es la segunda: aunque, en la práctica, podemos simular algunos procesos, el universo en sí no puede ser el resultado de una computación, porque esa computación requeriría una memoria y capacidad de proceso infinitas. Dado que cualquier parte del universo posee una cantidad ilimitada de información, nuestra mente, en principio, tampoco es simulable. Por supuesto, es posible encontrar buenas aproximaciones usando ordenadores, pero no se puede establecer una relación uno a uno entre sucesos en nuestra mente y operaciones en un ordenador. Este punto de vista está, sin embargo, en declive, gracias al avance de nuestra comprensión de la gravedad cuántica.

La mecánica cuántica y la relatividad general, que es la teoría relativista de la gravedad, son incompatibles. Mientras que la teoría de Einstein considera que el espacio es continuo e infinitamente divisible, en la mecánica cuántica la materia está dividida en pequeñas partes indivisibles llamadas *cuantos*. Esto produce un conflicto entre las dos teorías, que hacen predicciones contradictorias para sucesos donde existen enormes campos gravitatorios pero los efectos cuánticos no pueden ignorarse. En otras palabras, si aplicamos la filosofía cuántica al espacio-tiem-

po en sí, obtenemos un ente discontinuo y rugoso, no la suavidad predicha por Einstein.

Este conflicto entre ambas teorías llevó y sigue llevando de cabeza a los físicos, que han pasado casi cien años tratando de encontrar una forma de reconciliarlas. A pesar de que se han logrado grandes avances, seguimos lejos de una teoría contrastable que logre unir las de forma satisfactoria.

Uno de esos grandes avances fue el trabajo del célebre físico teórico y divulgador británico Stephen Hawking (n. 1942) sobre *agujeros negros*, esas estrellas cuya gravedad es tan potente que se colapsan sobre sí mismas, hasta tal punto que ni siquiera la luz es capaz de escapar de ellas. Según la relatividad general, nada puede escapar de un agujero negro: su interior está, a todos los efectos, desconectado del resto del universo. Hawking descubrió que los agujeros negros no son realmente negros, sino que tienen que evaporarse lentamente emitiendo radiación, como si fueran un cuerpo caliente. Eso le llevó a la idea de que los agujeros negros tienen que poseer algo llamado *entropía*.

La entropía es un concepto de la física estadística que suele asociarse al desorden. Sin embargo, una descripción más precisa es que la entropía nos da información sobre los posibles estados microscópicos de nuestro sistema. Si, por ejemplo, observamos que un gas tiene determinada presión, volumen y temperatura, sabemos que se encuentra en uno de la multitud de estados microscópicos posibles. Todos ellos darán lugar a las propiedades que estamos midiendo pero no tenemos forma alguna de saber en cuál de esos estados microscópicos se encuentra. La entropía nos dice cuántos de esos estados microscópicos existen: cuantos más haya, más probable es que observemos el sistema en esa configuración.

El concepto de entropía resulta estar íntimamente ligado a la computación, porque nos dice cuánta información contiene el sistema que estudiamos. Una forma de ver por qué es pensar que, cuantas más partículas tiene un gas y más volumen ocupan, más información requerimos para especificar la posición de todas y cada una de ellas. Dado que el gas «sabe» dónde se encuentran todas esas partículas, contiene toda esa información.

¿CUÁNTA INFORMACIÓN CABE EN UN AGUJERO NEGRO?

El resultado que demuestra que la energía contenida en un agujero negro es proporcional a su área puede obtenerse de forma relativamente sencilla, usando varias simplificaciones. En este caso, supondremos que cada partícula dentro de un agujero negro contiene una unidad de información; para saber cuánta información hay almacenada, intentaremos contar cuántas partículas puede haber, como máximo, en un agujero negro. Un agujero negro es un cuerpo en el que la *velocidad de escape*, mostrada en la figura 1, es decir, la velocidad que se necesita para escapar a su atracción gravitatoria, es igual a la de la luz. La velocidad de escape para un planeta con masa M y radio R viene dada por la siguiente fórmula:

$$v^2 = \frac{2GM}{R}$$

donde G es la constante de la gravitación universal. Para derivar el resultado de Hawking, reorganizaremos la ecuación para aislar la masa del agujero negro y sustituiremos la velocidad, v , por la velocidad de la luz, c obteniendo:

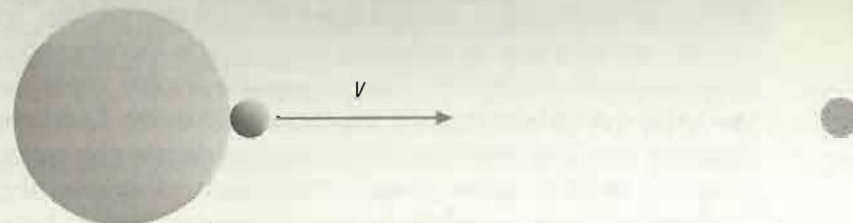
$$M = \frac{c^2 R}{2G}$$

De esta ecuación, lo único que nos interesa es que la masa de un agujero negro es proporcional a su radio. Dado que la masa es una forma concentrada de energía, tal como expresó Einstein en su famosa ecuación $E=mc^2$, podemos concluir que la energía almacenada en un agujero negro tiene que ser proporcional a su radio. El siguiente ingrediente que necesitamos es la mecánica cuántica. Nos preguntamos cuál es la mínima energía que puede tener un fotón, que es la partícula de la luz, en un agujero negro. En este caso, la energía es inversamente proporcional a su *longitud de onda*, que es la distancia entre dos crestas. Si deseamos encontrar la mínima energía de un fotón, tenemos que encontrar cuál es la máxima longitud de onda que puede tener. Esta tiene que ser el perímetro del agujero negro, como ilustra la figura 2. Dado que el perímetro del agujero negro es proporcional a su radio, tenemos que la mínima energía que puede tener un fotón es inversamente proporcional al radio del agujero negro. Matemáticamente:

$$E_{\text{mínima}} \propto \frac{1}{R}$$

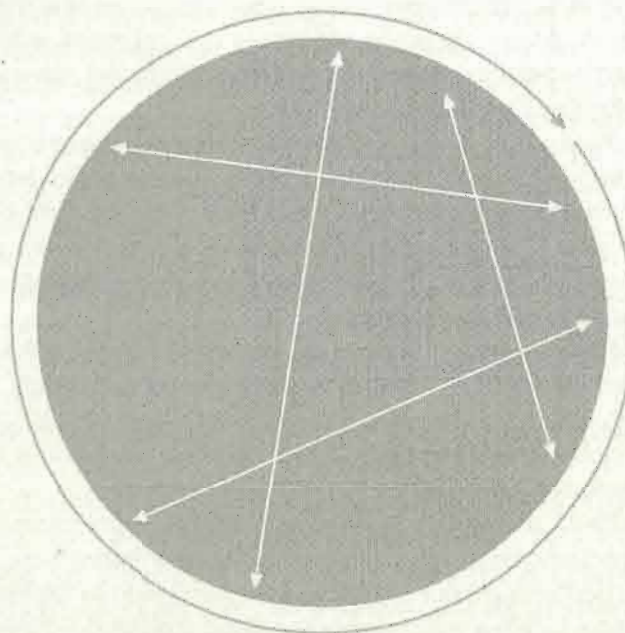
donde el símbolo \propto significa «es proporcional a». Si ahora queremos saber cuántos fotones caben en el agujero, tenemos que dividir su energía total por la energía mínima de un fotón. Dado que la energía total es proporcional al radio y la energía del fotón es inversamente proporcional a este, el número de fotones tiene que ser proporcional al radio al cuadrado. Pero el radio al cuadrado es proporcional al área del agujero negro: por lo tanto, la información del agujero negro es proporcional a su área.

FIG. 1



La velocidad de escape es la velocidad que debe tener un cuerpo en la superficie para escapar a la influencia gravitatoria del planeta o estrella donde se encuentre.

FIG. 2



El perímetro del agujero negro corresponde a la máxima longitud de onda posible. Cualquier otra distancia, representada por las dobles flechas, será menor.

El hecho de que los agujeros negros tengan entropía nos dice que contienen información; podríamos decir que son una especie de disco duro cósmico. Y no solo eso: un agujero negro contiene la máxima cantidad de información posible. Para almacenar un bit hace falta materia y, cuantos más bits queremos guardar, más materia necesitamos. Si queremos almacenar la mayor cantidad de información posible en el menor espacio posible, tendremos que hacer que nuestra materia sea densa en extremo. La máxima densidad que podemos obtener se da cuando nuestra materia colapsa y forma un agujero negro. Por lo tanto, un agujero negro contiene la mayor cantidad de información que puede almacenarse en una cierta región del espacio.

Es decir que, por un lado, los agujeros negros tienen entropía y por lo tanto información; por otro, la información que contienen es la máxima que puede almacenarse en una región del espacio. Ahora podemos utilizar lo que sabemos sobre agujeros negros para averiguar si una región finita del espacio contiene información infinita o limitada.

Para sorpresa general, Hawking descubrió que la información tiene un límite y, por si fuera poco, que ese límite viene dado por el área del agujero negro, no por su volumen. Eso significa que no solo la información contenida en una región del universo es finita, sino que es proporcional al área de su alrededor, al contrario de lo que cabría esperar.

Una revelación que parece alentar la idea de que el universo, al contener una cantidad de información delimitada, sí es computable. Por extensión, nuestra mente, que ocupa una región limitada del espacio-tiempo, tiene que ser computable también, al menos en principio.

En la actualidad, la idea de Hawking forma parte de una tendencia general a transformar problemas sobre sistemas físicos en problemas sobre información, siguiendo la famosa frase del físico teórico estadounidense John Archibald Wheeler (1911-2008) «it from bit», algo así como que de los bits procede todo; es decir, lo físico es producto de la información y no a la inversa. Hoy en día cada vez más físicos están convencidos de que la respuesta al encaje entre la mecánica cuántica y la relatividad

tiene que pasar por abandonar ideas anticuadas como el espacio y el tiempo y sustituirlas por otras leyes que se refieran en exclusiva a la información y su proceso, siendo el espacio-tiempo un subproducto de conceptos más fundamentales. Si ese programa llega a buen término, el debate sobre lo que una máquina puede o no simular tomará un cariz muy distinto al actual.

¿CIENCIA O CONSCIENCIA?

Como se ha visto, la consciencia es un problema tan difícil que su estudio requiere conceptos de áreas dispares, desde las matemáticas fundamentales a la física, pasando por la computación. A pesar de lo que se suele pensar, los avances en ciencia básica no solo nos dan una mejor perspectiva sobre el universo, sino que nos pueden ayudar a entendernos a nosotros mismos.

Quizá pueda parecer extraño que teoremas sobre la fundamentación de las matemáticas tengan relevancia alguna sobre el fenómeno de la consciencia, pero se trata de un hecho inevitable: cualquier sistema que pueda referirse a sí mismo es susceptible a un gran número de paradojas. Si consideramos que una persona es equivalente a la información contenida en su cerebro, veremos que entre esa información se encuentra un concepto, el yo, que se refiere a toda ella. En términos informáticos, tenemos una colección de símbolos abstractos, uno de los cuales se refiere a todo el conjunto, creando un efecto similar al de dos espejos paralelos, donde cada imagen contiene todas las anteriores.

Física y consciencia son inseparables, porque la física estudia las leyes que rigen el universo y las personas somos, nos guste o no, una parte inseparable de ese universo. No podemos separar entre observador y observado: las personas somos un ejemplo del universo observándose a sí mismo. Esto dificulta el análisis tanto del universo como de la consciencia: por un lado, los físicos no pueden ignorar el hecho de que forman parte de lo que estudian, a pesar de que les gustaría tener una visión lo más objetiva posible. Por otro, las personas no podemos ignorar que

formamos parte de un universo con leyes, aunque nos sea mucho más fácil fijarnos solo en nuestra intuición.

En esta situación circular, donde el observador se observa a sí mismo mientras intenta no influir en lo que observa, es una necesidad que surjan paradojas aparentes con las que hay que lidiar si se quiere tener una mejor comprensión de lo que pasa. Tras un período de confusión, las contradicciones se despejan para dar paso a una mejor comprensión del universo y de nosotros mismos. La ciencia avanza a base de encontrar contradicciones y despejarlas.

Entender la consciencia exige salir de nuestro ser para pensar sobre nosotros mismos: objetivar la subjetividad. En esta empresa, más que en cualquier otra, las respuestas que nos da la física atentan contra la intuición. Al final, tenemos dos opciones: o negamos lo que intuimos en favor de la ciencia o rechazamos la ciencia en pro de la intuición. Hagamos lo que hagamos, nuestra decisión determinará lo que el universo opina de sí mismo.

- DEAÑO, A., *Introducción a la lógica formal*, Madrid, Alianza Editorial, 2004.
- DEUTSCH, D., *La estructura de la realidad*, Barcelona, Anagrama, 2002.
- GREENE, B., *El universo elegante. Supercuerdas, dimensiones ocultas y la búsqueda de la teoría final*, Barcelona, Editorial Crítica, 2011.
- HOFSTADTER, D.R., *Gödel, Escher, Bach*, Barcelona, Booket, 2015.
- : *Yo soy un extraño bucle*, Tusquets Editores, 2008.
- NAGEL, E., NEWMAN, J.R., *El teorema de Gödel*, Madrid, Editorial Tecnos, 2007.
- PARFIT, D., *Razones y personas*, Madrid, A. Machado Libros, 2005.
- PENROSE, R., *La nueva mente del emperador*, Barcelona, Editorial Debolsillo, 2015.
- : *Las sombras de la mente. Hacia una comprensión científica de la consciencia*, Barcelona, Editorial Crítica, 2000.
- PINKER, S., *Cómo funciona la mente*, Barcelona, Destino, 2008.
- SMULLYAN, R., *¿Cómo se llama este libro? El enigma de Drácula y otros pasatiempos lógicos*, Madrid, Ediciones Cátedra, 2007.
- THORNE, K., *Agujeros negros y tiempo curvo*, Barcelona, Editorial Crítica, 2000.

- agujero negro 79, 149-152
 amplitud 24, 34, 35, 49
 Aristóteles 16, 17
 axioma 111, 131-133

 bit 127, 130, 137, 147, 152
 Bohr, Niels 49, 55
 Born, regla de 35, 42, 51
 borrador cuántico de elección
 retardada 55, 57

 caos determinista 89-91, 93,
 100
 Carroll, Sean 50
 Chalmers, David 123
 completitud 131, 132
 computable 137-140, 142, 145,
 146, 152
 consciencia 7-11, 13, 15, 19-21,
 50, 53-55, 58, 60, 61, 77, 80, 83,
 103, 105, 113, 115, 117-120,
 122-126, 131, 140, 153, 154

 consistencia 145
 contracción de longitudes 72,
 76, 77
 Conway, John H 110-113
 Copenhagen, interpretación de
 48-53, 60, 97, 100, 101, 103,
 108-110, 140
 cuántico, teletransporte 44, 104,
 105
 cuanto 148

 decoherencia 38, 39, 46, 47, 49,
 52, 101, 140
 Demócrito 17, 86
 Descartes, René 16, 18, 19
 determinismo 91, 108
 Deutsch, David 50, 53, 108, 109
 diagrama espacio-temporal 94,
 95, 98, 99
 dilatación temporal 66, 71, 72,
 76, 77, 79, 125
 dualismo 16, 18, 19

Einstein, Albert 30, 38, 50, 61, 66, 74, 75, 82, 96, 111, 126, 148-150
 electromagnetismo 63
 emulación 122
 entrelazamiento 40, 44, 45, 58
 entropía 149, 152
 espacio continuo 146, 148
 espacio-tiempo 70, 72, 74, 77-82, 91, 92, 95-97, 108, 109, 152, 153
 espín 40, 41, 44, 51, 111
 éter luminífero 63, 65, 67
 Everett, Hugh 50, 55
 experimento de la doble rendija 21, 25, 28, 31-34, 38, 39, 44, 46, 47, 100
 Faraday, Michael 63
 función de onda 35, 36, 38, 40, 41, 44, 47, 49-54, 58, 138, 140
 colapso de la 35, 38, 40, 41, 44, 47, 51, 53, 138, 140
 Galilei, Galileo 16-18
 Gödel, Kurt 131, 133, 134
 teorema de 131, 133, 134, 137, 139, 140, 145
 gödelización 134, 135
 habitación china, argumento de la 119
 Hameroff, Stuart 140, 141
 Hawking, Stephen 149, 150, 152
 Hofstadter, Douglas R. 117, 120-122
 Huxley, Thomas Henry 18-20
 identidad personal 104-106
 información 7, 36, 44, 47, 49, 51, 54-56, 58, 60, 81, 87, 106, 111, 112, 118, 127, 129, 144-150, 152, 153
 interferencia 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 38, 46, 47, 55, 56, 57, 58, 59, 60, 67
 intervalo
 espacial 77, 78, 94, 96, 98, 99
 temporal 68, 70, 72, 77, 78, 94, 96, 98
 Kochen, Simon B. 110-112
 libre albedrío 8-10, 17, 18, 83, 85-88, 90-92, 96, 97, 100, 101, 103-105, 107, 108, 110-114
 luz, velocidad de la 63-68, 70, 94, 95, 98, 111, 150
 Maxwell, James Clerk 63, 64
 medida, problema de la (incertidumbre) 13, 33, 38, 47, 88, 89, 97, 111, 112
 metafísica 112, 113, 120
 metateorema 137
 Michelson-Morley, experimento de 63, 65, 67
 microtúbulos 140, 141
 Minkowski, geometría de 92
 monismo 16-18, 20
 múltiples universos,
 interpretación de 50-54, 58, 97, 103, 104, 106, 108-110, 140
 multiverso 108, 109, 140
 Nagel, Thomas 120-123, 126
 Newton, Isaac 9, 16, 24, 34, 48, 86, 88, 89, 91, 96

número
 racional 146, 147
 irracional 146, 147
 observador 8, 33, 38, 39, 44, 47, 49-51, 53, 54, 56, 58, 60, 64-66, 68, 69-72, 76, 78, 80, 94, 98, 99, 153, 154
 operador 36, 42, 127, 128, 129
 parada, problema de 139
 Parfit, Derek 104
 Penrose, Roger 55, 118, 123, 126, 131, 137, 138, 140-143, 146
 Pinker, Steven 144
 poliminó 142, 143
 postulado 35, 63, 66, 68, 70, 96
 probabilidad, onda de 34, 35, 45
 puerta lógica 127, 130, 136, 137, 146
 red neuronal 141, 144, 145
 reduccionismo 9, 118, 120, 122-126
 regla de inferencia 131-133, 137
 relatividad especial 61, 66, 80, 92, 96, 97, 111, 138
 Schrödinger, Erwin 39
 ecuación de 35-37, 51, 103
 gato de 39, 47
 Searle, John 119, 120, 123, 125, 126
 sensibilidad a las condiciones iniciales 89
 simulación 117, 118, 146, 148
 subjetividad, argumento de la 8, 120, 121, 154
 Tegmark, Max 140
 teorema 71, 73, 77, 110-112, 131-135, 137, 139, 140, 147, 153
 tiempo 9, 11, 15, 20, 36, 56, 61, 66, 68-72, 74-82, 85, 87, 90-96, 98, 99, 104, 105, 114, 137, 145, 153
 Turing, Alan 118, 120, 126, 129, 139
 máquina de 126, 129, 131, 133, 137, 145
 test de 118-120
 Wheeler, John Archibald 152
 Wolfram, Stephen 96, 101
 Young, Thomas 24, 25, 31
 zombis filosóficos, argumento de los 123-125

Ciencia y consciencia

La física del siglo xx tuvo enormes consecuencias no solo en nuestra comprensión del universo, sino en la de la consciencia. La mecánica cuántica obligó a los físicos a reflexionar sobre el papel del observador en la ciencia, mientras que la relatividad puso de manifiesto el conflicto entre el tiempo físico y el psicológico. Las dos grandes teorías del siglo pasado nos invitan a reflexionar sobre la naturaleza de la consciencia y su papel en el devenir del universo.

Eduardo Arroyo es físico, escritor y divulgador científico.